



**UNIVERSIDADE DO SUL DE SANTA CATARINA
ERICKSON LEON KOVALSKI**

**NO UNIVERSO DOS BITS E CÉREBROS
UMA EXPLORAÇÃO DA TEORIA DA INFORMAÇÃO ALGORÍTMICA.**

Palhoça
2023

ERICKSON LEON KOVALSKI

**NO UNIVERSO DOS BITS E CÉREBROS
UMA EXPLORAÇÃO DA TEORIA DA INFORMAÇÃO ALGORÍTMICA**

Trabalho de conclusão de curso apresentado ao Curso de Graduação em Matemática, Universidade do Sul de Santa Catarina – UNISUL como parte dos requisitos da obtenção do grau de Bacharel.

Orientador: Prof. Me. Mário Selhorst

Palhoça
2023

ERICKSON LEON KOVALSKI

**NO UNIVERSO DOS BITS E CÉREBROS
UMA EXPLORAÇÃO DA TEORIA DA INFORMAÇÃO ALGORÍTMICA**

Trabalho de conclusão de curso apresentado ao Curso de Graduação em Matemática, Universidade do Sul de Santa Catarina – UNISUL como parte dos requisitos da obtenção do grau de Bacharel.

Palhoça, 26 de junho de 2023.

Professor Me. Mário Selhorst (Orientador)
Universidade do Sul de Santa Catarina – UNISUL

Professor Me. Dalmo Gomes de Carvalho
Universidade do Sul de Santa Catarina – UNISUL.

Professora Dr^a. Ana Regina de Aguiar Dutra
Universidade do Sul de Santa Catarina – UNISUL.

Podemos considerar o estado atual do universo como o efeito de seu passado e a causa de seu futuro. Uma inteligência que, em um determinado momento, conhecesse todas as forças que colocam a natureza em movimento e todas as posições de todos os itens dos quais a natureza é composta, se essa inteligência também fosse vasta o suficiente para submeter esses dados à análise, ela abarcaria em uma única fórmula os movimentos dos maiores corpos do universo e aqueles do menor átomo; para tal inteligência nada seria incerto e o futuro, assim como o passado, estaria presente diante de seus olhos.

— Laplace

RESUMO

O objetivo deste Trabalho de Conclusão de Curso é realizar um estudo introdutório e abrangente sobre a Teoria da Informação Algorítmica (AIT), evidenciando suas diversas aplicações em variados domínios de pesquisa. A metodologia empregada consiste em uma revisão bibliográfica conceitual, baseada principalmente em artigos científicos. Se baseando no conceito de complexidade, a pesquisa explora as contribuições de diversos cientistas, como Claude Shannon, Alan Turing e Norbert Wiener, que foram fundamentais para o avanço do entendimento de informação, computação e sistemas complexos. O texto analisa a evolução da produção científica na AIT, destacando os marcos significativos, bem como os temas e áreas de pesquisa emergentes. Foca, especialmente, nas contribuições dos pioneiros da AIT, como Ray Solomonoff, Andrey Kolmogorov e Gregory Chaitin. Conceitos-chave como entropia e máquinas de Turing são explorados para investigar a natureza probabilística da informação e a capacidade dos sistemas de processá-la. Por fim, com o uso de uma linguagem acessível e exemplos ilustrativos, sem negligenciar o rigor matemático e computacional, o estudo busca contribuir para a disseminação e compreensão da AIT entre um público mais amplo. Em conclusão, a presente pesquisa reforça a relevância e influência da AIT na nossa compreensão dos sistemas complexos e sugere direções para estudos futuros.

Palavras-chave: Complexidade, Teoria Algorítmica da Informação, Aleatoriedade, Algoritmos, Máquina de Turing.

ABSTRACT

This final project assignment provides an introductory and comprehensive examination of Algorithmic Information Theory (AIT), illustrating its myriad applications across various research domains. Our methodology comprised a bibliographic review, primarily of scientific articles. The research is rooted in the concept of complexity and sheds light on the seminal contributions of scientists such as Claude Shannon, Alan Turing, and Norbert Wiener, whose work advanced our understanding of information, computing, and complex systems. We traced the evolution of scientific discourse in AIT, pinpointing pivotal milestones and emerging research themes. Special emphasis was placed on the pioneering work in AIT by figures such as Ray Solomonoff, Andrey Kolmogorov, and Gregory Chaitin. The study scrutinizes key concepts like entropy and Turing machines to probe the probabilistic nature of information and the capacity of systems to process it. Using accessible language and clear examples, without compromising on mathematical and computational rigor, we aimed to foster a wider understanding and dissemination of AIT. In conclusion, the present research underscores the relevance and influence of AIT in our understanding of complex systems and suggests directions for future studies.

Key words: Complexity, Algorithmic Information Theory, Randomness, Algorithms, Turing Machines.

LISTA DE FIGURAS

FIGURA 1: REPRESENTAÇÃO ESQUEMÁTICA DE UM TERMOSTATO, RESSALTANDO AS SETAS DE TROCA DE INFORMAÇÃO (INPUT) E CONTROLE (OUTPUT). FONTE: ELABORAÇÃO DO AUTOR	31
FIGURA 2: TRANSIÇÕES DE ESTADOS E PROBABILIDADES ASSOCIADAS. FONTE: ELABORAÇÃO DO AUTOR.	40
FIGURA 3: EXEMPLO DE OPERAÇÕES DE REALIZADAS POR UMA TM AO INCREMENTAR O NÚMERO BINÁRIO 1011 EM UMA UNIDADE. FONTE: ELABORAÇÃO DO AUTOR.	48
FIGURA 4: IMAGEM DE UMA TORRE DE LEGOS COM PADRÃO MAIS COMPLEXO DE CORES GERADO ATRAVÉS DA TECNOLOGIA <i>STABLE DIFFUSION</i> PELO <i>PROMPT</i> : A LEGO TOWER OF A FEW BLOCKS OF ALTERNATING COLORS.	58

LISTA DE QUADROS

QUADRO 1: PROCESSO DE OBTENÇÃO DA EQUAÇÃO DO MAPA LOGÍSTICO. FONTE: ELABORAÇÃO DO AUTOR.....	23
QUADRO 2: SÉRIES DE APROXIMAÇÕES AO INGLÊS COM LETRAS EXTRAÍDOS DE SHANNON(1948), P.7.	44
QUADRO 3: IMPLEMENTAÇÃO EM PSEUDOCÓDIGO DO QUE SERIA UM PROGRAMA QUE AVALIA SE O PROGRAMA P COM A ENTRADA I TERMINA OU NÃO. AO FINAL O PROGRAMA TOMA A SI MESMO COMO OBJETO. FONTE: ELABORAÇÃO DO AUTOR.....	50

LISTA DE TABELAS

TABELA 1: OPERAÇÕES DE UMA TM. FONTE: ELABORAÇÃO DO AUTOR.	47
TABELA 2: TABELA DE INSTRUÇÕES PARA ADIÇÃO BINÁRIA, EQUIVALENTE AO QUE HOJE CHAMARÍAMOS DE UM PROGRAMA. FONTE: ELABORAÇÃO DO AUTOR.....	48

LISTA DE ABREVIATURAS E SIGLAS

Abreviação	Significado	Tradução para o Português (se terminologia estrangeira)
AEC	Antes da Era Comum	
AIT	Algorithmic Information Theory	Teoria da Informação Algorítmica
EC	Era Comum	
KC	Kolmogorov Complexity	Complexidade de Kolmogorov
MDL	Minimum Description Length	Descrição de Comprimento Mínimo
TM	Turing Machine	Máquina de Turing
UTM	Universal Turing Machine	Máquina de Turing Universal

SUMÁRIO

1	INTRODUÇÃO.....	12
2	COMPLEXIDADE.....	18
2.1	Antes do bit veio o contínuo.....	19
2.2	Mr. Maxwell tem um demônio.....	25
2.3	Ressonâncias rítmicas: Cibernética.....	29
2.4	Átomos de Informação.....	32
3	TEORIA DA INFORMAÇÃO.....	38
3.1	O caminho menos percorrido.....	39
3.2	Surpresas na transmissão.....	42
3.3	Máquinas Inteligentes: Turing.....	47
4	TEORIA ALGORÍTMICA DA INFORMAÇÃO.....	52
4.1	Problema da indução em Solomonoff.....	53
4.2	Complexidade de Kolmogorov.....	58
4.3	Descrição de Comprimento Mínimo.....	60
4.4	Número ômega.....	64
5	REFLEXÕES FINAIS.....	66
	REFERÊNCIAS.....	71
	APÊNDICE A: CÓDIGO FONTE DOS EXEMPLOS E VISUALIZAÇÕES UTILIZADOS NESTE TRABALHO.....	75
	APÊNDICE B: PUBLICAÇÕES EM TEORIA ALGORÍTMICA DA INFORMAÇÃO.....	83

1 INTRODUÇÃO

A verdade está nua; mas debaixo da nudez está o desvelado. O que é simples é sempre falso. O que não é simples, não tem utilidade.
Paul Valéry

Complexidade é um conceito multifacetado que se revela nas mais diversas áreas de pesquisa que abrangem desde a Biologia e a Física até as Ciências Sociais e a Ciência da Computação. Em sua essência, a complexidade refere-se à natureza intrincada e interdependente dos sistemas, caracterizada por numerosos elementos, interações e propriedades emergentes. Essas propriedades não podem ser facilmente compreendidas examinando os componentes individuais isoladamente. O estudo da complexidade, portanto, procura desvendar esses padrões subjacentes, as dinâmicas e comportamentos que emergem das interações dentro de sistemas complexos. Sistemas estes que podem variar desde a escala microscópica de partículas subatômicas até as dimensões do universo, incluindo ecossistemas, economias globais e mesmo o cérebro humano.

A complexidade é frequentemente fruto de relações não lineares, ciclos de *feedback* e do comportamento coletivo dos componentes individuais, engendrando fenômenos surpreendentes e, por vezes, paradoxais. Um exemplo é o comportamento de um bando de pássaros ou de um cardume de peixes. Observando de fora, esses padrões de comportamento coletivo, como movimentos sincronizados ou reviravoltas abruptas, podem sugerir a existência de uma autoridade central ou de um líder determinado que direciona o grupo inteiro. Porém, nesses sistemas, cada ave ou peixe individual segue regras simples baseadas em seu entorno imediato, como manter uma distância mínima dos vizinhos ou alinhar sua direção com os indivíduos próximos. Estas interações locais dão origem a padrões emergentes ao nível de grupo, onde o comportamento coletivo emerge sem nenhum controle central. Essa auto-organização permite que o grupo responda rapidamente às mudanças no ambiente e exiba movimentos complexos e coordenados.

Uma característica típica dos sistemas complexos é a sua dependência sensível às condições iniciais e a possibilidade de desencadear efeitos em cascata. Alterações mínimas nas condições de partida ou nos parâmetros que orientam as

interações podem culminar em desvios consideráveis no comportamento do sistema ao longo do tempo. Essa característica, que recebeu o apelido de *efeito borboleta*¹, sublinha a imprevisibilidade e a natureza não determinista dos sistemas complexos.

Para sua compreensão e modelagem, a complexidade requer uma abordagem interdisciplinar que envolve a incorporação de conceitos e metodologias de vários campos de estudo: A Matemática, a Física e a Ciência da Computação desempenham um papel fundamental nesse processo, fornecendo estruturas conceituais e ferramentas analíticas para examinar e entender sistemas complexos; a Biologia, por sua vez, revela a complexidade através da observação de redes intrincadas de interações dentro dos ecossistemas, na auto-organização de células e organismos, bem como nos processos evolutivos que moldam a biodiversidade; as Ciências Sociais abordam a complexidade inerente às sociedades humanas, que inclui a análise de redes sociais, a investigação de dinâmicas culturais e a compreensão de como comportamentos coletivos emergem dentro desses sistemas sociais; por fim, a Economia oferece pontos de vista sobre como os mercados evoluem, de como os mecanismos dos sistemas financeiros se relacionam e de como as redes econômicas complexas se comportam.

Não se limitando a essas áreas, a complexidade reside no centro da pesquisa científica. O reducionismo é uma ferramenta indispensável para pesquisa, porém é preciso também reconhecer que ele é inerentemente limitado. E quando queremos, enquanto cientistas, expandir nosso horizonte de compreensão, é evidente a falta de uma linguagem para expressar essas preocupações. Historicamente, o fazer científico se tornou muito eficiente em explorar campos cada vez mais estreitos. O progresso é incremental. Porém, quando o objetivo é conectar os diversos saberes, nossas ferramentas são limitadas e acabamos com uma visão míope do mundo, centrada no espaço que, por acaso, ocorre de estarmos.

Desde os anos sessentas do século XX, mais e mais essas barreiras entre as ciências particulares se mostram demasiado estreitas para dar conta dos desafios contemporâneos. Como discutir questões contenciosas como as mudanças

¹ Esse apelido vem do título de uma apresentação proferida pelo meteorologista Edward Lorenz, um dos fundadores da pesquisa em sistemas caóticos: *Predictability: does the flap of a butterfly's wings in Brazil set off a tornado in Texas?* (LORENZ, 1972)

climáticas sem um protocolo científico de pesquisa? Como relacionar variáveis como a ascensão de ideologias específicas com mudanças no cenário econômico, com o surgimento de tecnologias e com transformações geopolíticas sem confundir causalidade com correlação? Ou, como planejar para o futuro, desde o ponto de vista do indivíduo até a escolha de políticas globais, sem uma forma de avaliar a relevância que dados oriundos de pesquisas acadêmicas possuem em determinado contexto?

Uma via eficaz para se obter compreensão de um fenômeno é através da criação de um modelo matemático análogo, que permita ao cientista interagir com o fenômeno simulado e testar hipóteses. Isso se aplica a uma ampla gama de fenômenos, desde a gravitação universal até a cognição humana, desde a dinâmica de interação entre predador e presa às flutuações do mercado financeiro.

Entender os limites dos modelos e de nossa capacidade de modelar é, mais que trazer sobriedade ao fazer científico, reconhecer a necessidade de se desenvolverem novos métodos para contornar as limitações conhecidas assim como abraçar novas perspectivas que desloquem posições estagnadas nos mais diversos setores da sociedade.

A matemática, intrinsecamente ligada à revolução tecnológica atual, permeia diversos setores da sociedade. De um lado, temos o avanço das tecnologias de computação, um progresso evidenciado na inovação de hardware, na expansão de data centers, e na crescente portabilidade e conectividade computacional. Por outro lado, temos o desenvolvimento de metodologias expressas em técnicas estatísticas de análise, algoritmos de compressão, de segurança e de comunicação de dados.

Com isto dito, introduzimos a perspectiva da Teoria Algorítmica da Informação (AIT) como um meio de navegar a complexidade. Esta teoria, surgindo no mesmo contexto histórico-cultural que outras interpretações da complexidade discutidas aqui, oferece uma visão de mundo que, a nosso ver, é extremamente frutífera para compreender fenômenos que habitualmente resistem aos métodos tradicionais.

O tema amplo deste trabalho concentra-se no campo multidisciplinar de Sistemas Complexos, que analisa sistemas com numerosos componentes interativos, tais como sistemas físicos, biológicos e sociais. O foco está na exploração dos recursos desses sistemas, como não-linearidade, loops de feedback, emergência e

auto-organização. No entanto, devido à vastidão do campo, o estudo delimita-se à aplicação da AIT nesses sistemas, enfocando as contribuições dos matemáticos Gregory Chaitin, Andrey Kolmogorov e Ray Solomonoff.

Um desafio significativo na área de Sistemas Complexos é entender e interpretar as informações geradas por esses sistemas. A AIT é uma ferramenta valiosa para este propósito. No entanto, sua compreensão pode ser trabalhosa, dificultando sua ampla aplicabilidade. O problema que esta pesquisa pretende abordar é, portanto, duplo: explorar como a AIT pode ser efetivamente aplicada no estudo de sistemas complexos e como tornar esta teoria acessível para uma compreensão mais ampla.

O objetivo geral desta pesquisa é, portanto, examinar a aplicação da AIT no estudo de Sistemas Complexos. Especificamente, busca-se não só entender melhor a questão que a AIT busca resolver, sua relevância para o avanço científico e suas diversas aplicações em diferentes domínios de pesquisa, mas também traduzir e simplificar esses conceitos para um público de língua portuguesa.

O presente trabalho se caracterizou como uma pesquisa bibliográfica de caráter qualitativo, cujo intuito foi explorar e analisar a literatura já publicada acerca do tema de interesse. Com o propósito de lançar as bases para pesquisas aprofundadas futuras, nossa abordagem primária foi baseada na revisão conceitual.

As fontes primárias consultadas consistiram em artigos dos principais autores pesquisados – Gregory Chaitin, Andrey Kolmogorov e Ray Solomonoff, obtidos por meio do Portal de Periódicos da CAPES. Além disso, acessamos os livros desses autores existentes em formato digital, seus textos disponibilizados em páginas de universidades ou livros físicos, por meio de bibliotecas universitárias.

Adicionalmente, foram consultadas as fontes secundárias através de uma pesquisa com termos relevantes. Para isto, foi utilizado o resultado de uma pesquisa preliminar como ponto de partida (descrita no Apêndice B). Esta etapa deu ênfase a seleção de trabalhos teóricos que realizam a discussão dos conceitos fundamentais da AIT que requeriam aprofundamentos a fim de serem apresentados no presente trabalho. Os artigos obtidos foram organizados com o auxílio da plataforma *Mendeley (Elsevier)* que além de ser responsável pelo controle da bibliografia utilizada, auxiliou na indexação dos conteúdos. Livros e textos adicionais, no entanto,

foram utilizados no decorrer do trabalho quando necessário para oferecer contextualização histórica ao surgimento dos conceitos.

O capítulo intitulado *Complexidade* abre nossa pesquisa e busca oferecer um panorama abrangente do advento de uma nova classe de problemas que as ferramentas convencionais se mostravam incapazes de abordar de maneira adequada. Estes problemas, considerados anomalias no contexto da *ciência normal*, começaram a sinalizar a necessidade de um novo paradigma, para usar a terminologia de Thomas Kuhn (PINGUELLI, 2008).

Essas anomalias, ao resistirem à interpretação e solução através dos métodos existentes, sugeriram a necessidade de uma reformulação no modo como entendemos e nos aproximamos desses problemas. Isso nos leva a um novo paradigma, que reconhece a existência de complexidade intrínseca em muitos sistemas e busca desenvolver novas ferramentas e abordagens capazes de lidar com ela. Este capítulo procura explorar essas ideias, estabelecendo as bases para a compreensão da natureza desses problemas complexos e da necessidade de inovação em nossa abordagem para resolvê-los.

No Capítulo 3, voltado aos primórdios da *Teoria da Informação*, prosseguimos com a construção da fundamentação teórica, apresentando duas barreiras para a compreensão de fenômenos complexos. Existem problemas que, em teoria, são solucionáveis, mas que na prática se mostram inacessíveis devido à demanda excessiva de recursos, como é o caso do problema dos três corpos. Há ainda problemas que são, por definição, intratáveis, entre os quais se inclui o problema da parada. Contudo, dispomos de ferramentas oriundas do campo da estatística que podem auxiliar na gestão de ambos os tipos de desafios.

O cerne deste trabalho reside no Capítulo 4, dedicado à AIT. Nele, temos a intenção de proporcionar uma visão abrangente dos conceitos fundamentais da teoria, os quais se encontram mais presentes na bibliografia específica.

No capítulo final, exploramos as relações entre a evolução biológica e o desenvolvimento das capacidades cognitivas como tópicos que podem ser abordados dentro do *framework* da AIT. Investigamos como os processos evolutivos moldam a cognição e examinamos como a AIT oferece uma perspectiva única para compreender a interação entre a evolução e a cognição. Essa análise visa ampliar nosso entendimento sobre a origem e a natureza das capacidades cognitivas nos

organismos vivos e, ao mesmo tempo, defender o uso sistemático da AIT como uma ferramenta de pesquisa nesta área interdisciplinar.

Desta forma, gostaríamos de enfatizar a relevância da AIT no cenário atual da ciência. Este trabalho tenta lançar luz sobre os conceitos mais importantes da teoria e suas implicações filosóficas, bem como discutir a aplicação prática dessa teoria nas pesquisas atuais. Com o aprofundamento de seu estudo, acreditamos que novos horizontes de conhecimento podem se abrir, oferecendo novas ferramentas para lidar com os desafios cada vez mais complexos que a ciência enfrenta.

2 COMPLEXIDADE

É interessante contemplar uma margem entrelaçada, coberta por muitas plantas de diversos tipos, com pássaros cantando nas moitas, com vários insetos voando e com minhocas rastejando pela terra úmida, e refletir que essas formas elaboradamente construídas – tão diferentes umas das outras e dependentes umas das outras de maneira tão complexa – foram todas produzidas por leis que atuam ao nosso redor.

Charles Darwin

Neste capítulo apresentamos uma análise de temas diversos que se entrelaçam em seus impactos e implicações para a compreensão dos sistemas complexos e caóticos. Com a contribuição de pesquisadores como Henri Poincaré, James Clerk Maxwell, Ludwig Eduard Boltzmann, Alfred J. Lotka, Vito Volterra, Richard M. Goodwin, Norbert Wiener, Walter Pitts e Warren McCulloch, exploramos os fundamentos teóricos da complexidade, caos, termodinâmica, biologia, economia, cibernética, teoria da informação e ciência cognitiva.

A discussão começa com a análise das descobertas de Poincaré sobre a natureza intrincada dos movimentos irregulares em sistemas mecânicos e a imprevisibilidade decorrente. Também consideramos as contribuições para a teoria do caos feitas por meio da Equação Logística e da constante de Feigenbaum. Em seguida, discutimos as noções de entropia e o desafio lançado pelo experimento mental do demônio de Maxwell à segunda lei da termodinâmica. Boltzmann e seus métodos estatísticos, ao desvendar a relação entre entropia e probabilidade, também ocupam um espaço significativo nesta seção. O terceiro tema abordado é a aplicação das equações de Lotka-Volterra para compreender as dinâmicas de sistemas biológicos, bem como a sua extensão para a análise de ciclos econômicos por Goodwin. A partir disso, adentramos no território da cibernética, explorando o conceito de feedback e a robustez dos sistemas biológicos frente ao ruído.

Finalmente, abordamos o mundo da informação e da computação, examinando a relevância dos bits como unidades fundamentais de informação e discutimos a estimativa de sinal e redução de ruído pelo filtro de Wiener. Além disso,

avaliamos as contribuições de Pitts, McCulloch, Von Neumann e Rosenblatt para o entendimento do sistema nervoso e do processamento de informação, bem como para a programação de computadores e aprimoramento do aprendizado das máquinas.

Ao longo do capítulo buscamos tecer as conexões entre essas disciplinas e campos de estudo, demonstrando como todos esses desenvolvimentos teóricos e descobertas têm moldado nossa compreensão dos sistemas complexos e caóticos, e pavimentando o caminho para as inovações tecnológicas e científicas.

2.1 Antes do bit veio o contínuo.

O estudo dos sistemas complexos e caóticos revolucionou nossa compreensão do mundo ao revelar a imprevisibilidade e a complexidade inerentes a muitos fenômenos naturais. No campo da mecânica, o chamado Problema dos Três Corpos impulsionou o desenvolvimento conceitual que deu origem à Teoria do Caos e ao campo de Sistemas Complexos ao abordar o problema aparentemente simples de determinar o movimento de três corpos celestes sob a influência da gravidade.

Desde Newton (1642 - 1726/27), houve avanços impressionantes no desenvolvimento de métodos analíticos para se resolverem problemas clássicos da mecânica. Porém, essas abordagens apresentam limitações quando lidam com uma ampla gama de problemas físicos. Computadores eletrônicos permitiram então aplicar métodos numéricos, que se tornaram populares tanto na ciência quanto na engenharia, sendo amplamente empregados em problemas práticos.

$$\begin{aligned}\ddot{\mathbf{r}}_1 &= -Gm_2 \frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} - Gm_3 \frac{\mathbf{r}_1 - \mathbf{r}_3}{|\mathbf{r}_1 - \mathbf{r}_3|^3} \\ \ddot{\mathbf{r}}_2 &= -Gm_3 \frac{\mathbf{r}_2 - \mathbf{r}_3}{|\mathbf{r}_2 - \mathbf{r}_3|^3} - Gm_1 \frac{\mathbf{r}_2 - \mathbf{r}_1}{|\mathbf{r}_1 - \mathbf{r}_2|^3} \\ \ddot{\mathbf{r}}_3 &= -Gm_1 \frac{\mathbf{r}_3 - \mathbf{r}_1}{|\mathbf{r}_1 - \mathbf{r}_3|^3} - Gm_2 \frac{\mathbf{r}_3 - \mathbf{r}_2}{|\mathbf{r}_2 - \mathbf{r}_3|^3}\end{aligned}$$

Eq. 1: Equações newtonianas do movimento para o vetor $\mathbf{r}_i = (x_i, y_i, z_i)$ com massas m_i . G é a constante gravitacional. Fonte: adaptado de HAND, FINCH (1998).

Em sistemas mecânicos integráveis, que são aqueles em que é possível encontrar soluções analíticas para as equações de movimento, o número de constantes de movimento é igual ao número de graus de liberdade do sistema. Isso significa que, para um sistema com n graus de liberdade, existem n constantes de

movimento (Eq. 1). No entanto, é importante ressaltar que essa é apenas uma condição necessária, pois também deve haver certa compatibilidade entre essas constantes, devido a razões geométricas (HAND, FINCH, 1998). Embora os sistemas integráveis representem apenas uma pequena fração dos sistemas possíveis, eles dominaram o campo de estudo da mecânica clássica, como exemplificado pelo movimento de dois corpos sob interação gravitacional mútua. No entanto, com a adição de um terceiro corpo, o problema se torna muito mais complexo (PINGUELLI, 2008). A descoberta do movimento caótico na dinâmica clássica por Henri Poincaré (1854–1912), no que está conhecido como Problema dos Três Corpos, foi decorrente da grande complexidade observada no movimento irregular na região de um ponto fixo hiperbólico (HAND, FINCH, 1998).

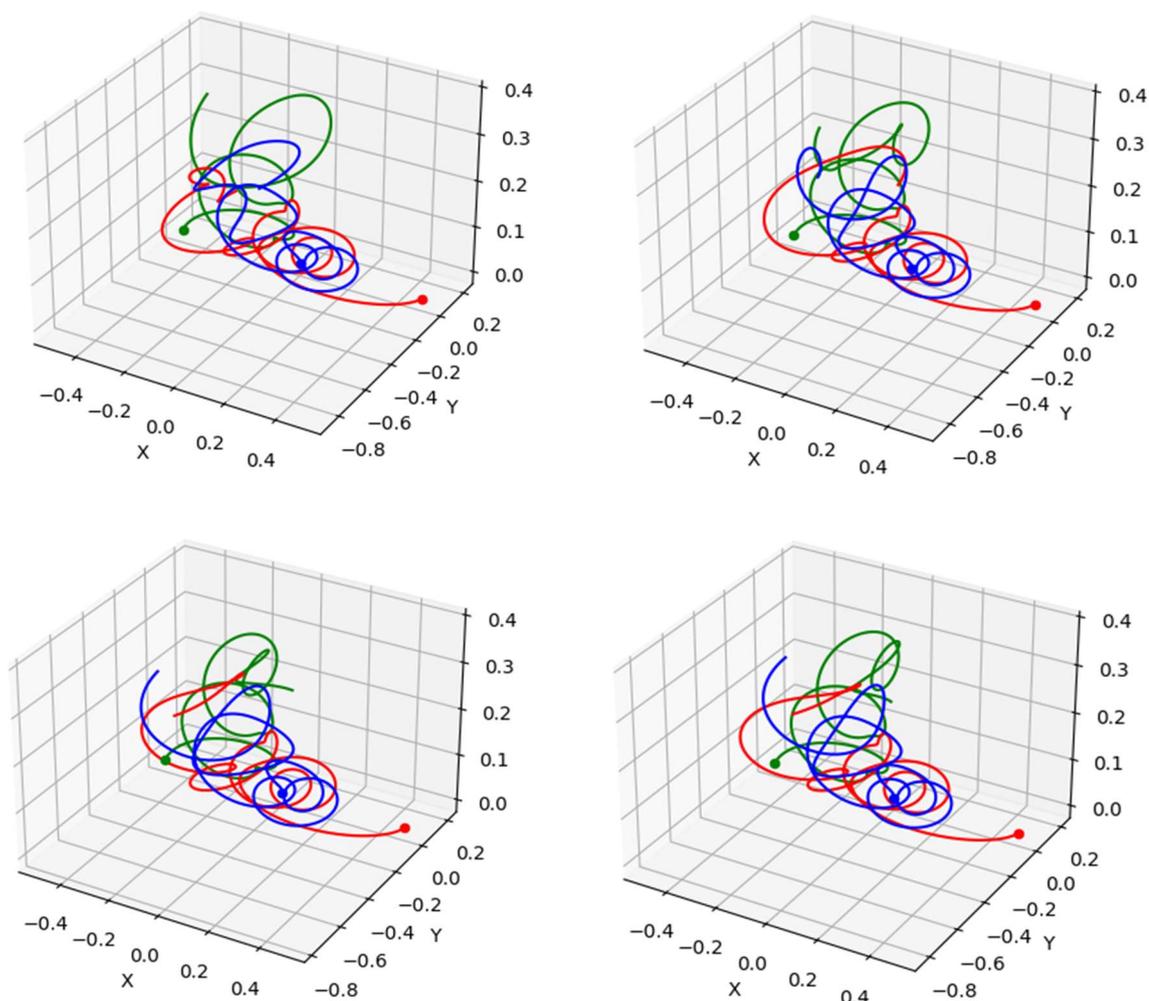


Gráfico 1: Quatro simulações do problema de três corpos em que pequenas mudanças nas condições iniciais levaram a diferentes desenvolvimentos do sistema. Neste caso a velocidade inicial do objeto em verde foi alterado em 0,001 a cada passo. Código disponível no Apêndice A. Elaboração do autor.

Em outras palavras, um ponto fixo hiperbólico é um ponto no espaço de fase do sistema em que as trajetórias próximas ao redor desse ponto divergem exponencialmente ao longo do tempo de forma que pequenas mudanças nas condições iniciais ou nos parâmetros do sistema podem resultar em comportamentos completamente diferentes no longo prazo (STONE, LEIGH, 2019). A presença de um ponto fixo hiperbólico indica uma instabilidade inerente ao sistema (Gráfico 1). Embora o comportamento seja governado pelas leis físicas e equações de movimento – portanto, determinístico –, pequenas perturbações podem levar a desvios grandes e imprevisíveis no comportamento das trajetórias das partículas ou componentes do sistema (LORENZ, 1972).

Outra elaboração neste sentido é a Equação Logística (Eq. 2), que é um modelo matemático fundamental o qual exemplifica muitos aspectos da complexidade em sistemas dinâmicos. Ela foi originalmente concebida por Pierre François Verhust (1804 – 1849) para descrever o crescimento populacional sob condições de recursos limitados (BACAËR, 2011).

A equação logística modela o crescimento populacional como uma função do tempo, onde a taxa de crescimento diminui à medida que a população se aproxima da capacidade de carga do ambiente. Quando a população é pequena em relação à capacidade de carga ($P \ll K$), a população cresce exponencialmente, pois há suficiente recurso disponível. No entanto, à medida que a população se aproxima da capacidade de carga ($P \cong K$), os recursos se tornam limitados, e a taxa de crescimento diminui até que a população se estabilize em K .

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K} \right)$$

Eq. 2: Equação logística. Fonte: BACAËR (2011)

Onde:

- P representa a população;
- t representa o tempo;
- r representa a taxa de crescimento intrínseco da população; e
- K representa a capacidade de carga do ambiente, que é o tamanho máximo da população que o ambiente pode sustentar a longo prazo.

Esta equação teve um papel influente no desenvolvimento da teoria do caos (GLEICK, 1987), um ramo da física que se dedica à análise de sistemas dinâmicos que exibem um comportamento sensível às condições iniciais. Quando submetida a certas modificações (como mostrado no Quadro 1), a equação logística pode produzir um comportamento dinâmico extremamente complexo, que inclui fenômenos como bifurcações (pontos onde a solução do sistema se divide) e atratores estranhos (conjuntos de condições iniciais que levam a trajetórias que se repetem de maneira não trivial), como ilustrado nos Gráficos 3 e 4. O diagrama de bifurcação do mapa logístico fornece uma representação visual do comportamento do sistema à medida que o parâmetro de taxa de crescimento², r , muda. Isso nos ajuda a compreender a transição de uma dinâmica simples (como equilíbrio ou oscilações periódicas) para um comportamento complexo e caótico.

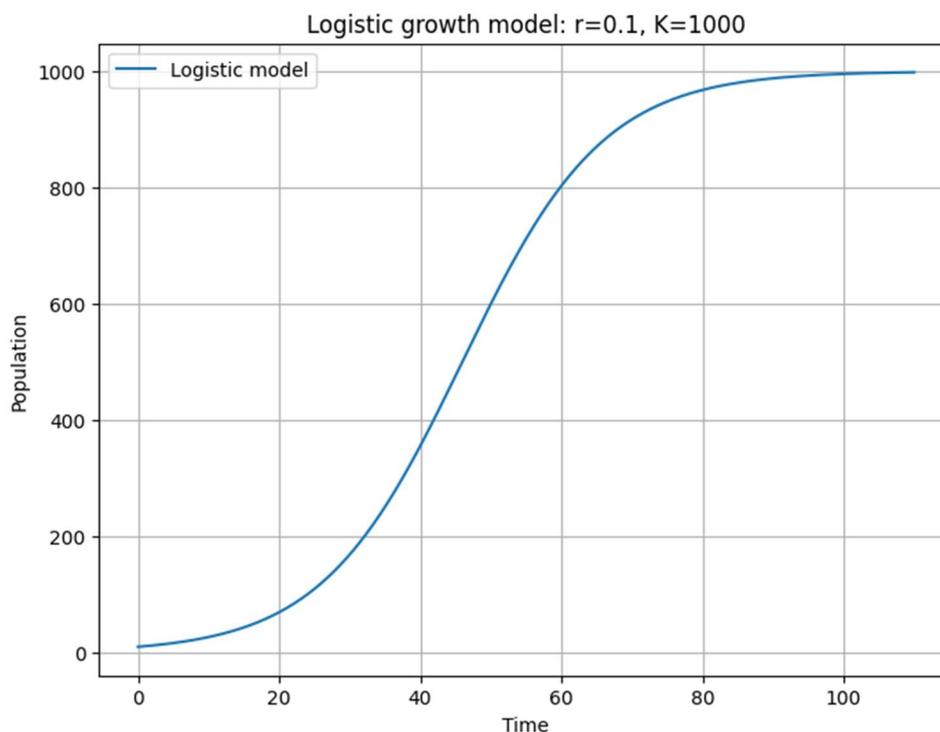


Gráfico 2: A função logística gera uma curva sigmóide que representa a evolução de uma população no decorrer do tempo. Código disponível no Apêndice A. Elaboração do autor.

² Também pode ser descrito como a taxa de fertilidade da população.

Em valores baixos de r , a população se estabiliza em um ponto fixo estável. Conforme r aumenta, observamos bifurcações: a população alterna entre dois valores, depois quatro, depois oito, e assim por diante. Isso representa a *ordem* presente no sistema. No entanto, em certos valores críticos de r , o comportamento se torna caótico – a população salta imprevisivelmente entre valores, nunca se estabelecendo em um padrão repetitivo. Apesar dessa aparente aleatoriedade, há uma quantidade surpreendente de estrutura nesse caos. Ao observar o diagrama, cópias menores do diagrama completo se repetem dentro das regiões caóticas. Isso é um fenômeno conhecido como *auto-similaridade* ou *fractalidade* e é uma característica comum dos sistemas caóticos.

Quadro 1: Processo de obtenção da equação do mapa logístico. Fonte: elaboração do autor

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K} \right)$$

$$\frac{d(xK)}{dt} = rxK \left(1 - \frac{xK}{K} \right)$$

$$\frac{d(x)}{dt} = rx(1-x)$$

Assumindo que o crescimento ocorre em intervalos discretos de tempo, vamos considerar t como uma unidade de tempo discreta, tal como 1 ano, $\Delta t = 1$, então Δx se torna a mudança na população por ano. Assim podemos calcular a população do ano seguinte x_{n+1} a partir da população atual x_n :

$$x_{n+1} = rx_n(1-x_n)$$

onde:

- x_n é o estado do sistema no tempo n ;
- x_{n+1} é o estado do sistema no próximo intervalo de tempo;
- r é um parâmetro que controla a dinâmica do sistema.

Essas bifurcações ocorrem em valores específicos de r , onde mudanças sutis no parâmetro resultam em mudanças drásticas no comportamento do sistema. É interessante notar que as diferenças entre esses valores de r onde ocorrem as bifurcações seguem uma relação importante: elas tendem a se aproximar de uma constante conhecida como constante de Feigenbaum (MITCHELL, 2009). Mitchell Jay Feigenbaum (1944 – 2019) foi um pioneiro nos estudos da teoria do caos e foi o primeiro a notar um padrão entre as bifurcações do mapa logístico em 1975.

Seja r_n o valor de r em que o sistema bifurca e passa a ter 2^n pontos de equilíbrio. Quando n é grande, podemos calcular a constante de Feigenbaum δ aproximadamente como a razão entre as diferenças entre os valores de r_n (Eq.3).

$$r_n = \frac{r_{n-1} - r_{n-2}}{r_n - r_{n-1}}$$

Eq. 3: Constante de Feigenbaum. Fonte: (MITCHELL, 2009)

Dessa forma, se observarmos os valores de r onde ocorrem as bifurcações (ou seja, onde o número de pontos de equilíbrio dobra), as razões entre as diferenças desses valores consecutivos de r tendem a se aproximar da constante de Feigenbaum (aproximadamente 4,6692016). No entanto, em um determinado valor crítico de r (em nosso caso, de aproximadamente 3,57), as bifurcações se tornam infinitas e o sistema entra em um regime caótico. A transição para o caos não significa que toda ordem seja perdida. Mesmo no regime caótico, existem certos valores para r que o sistema exhibe comportamento periódico, conhecidos como janelas periódicas. Essa interação entre o comportamento periódico e o caótico é um dos aspectos mais interessantes dos sistemas dinâmicos não lineares: ele não é aleatório. Os sistemas caóticos seguem regras determinísticas e seu comportamento,

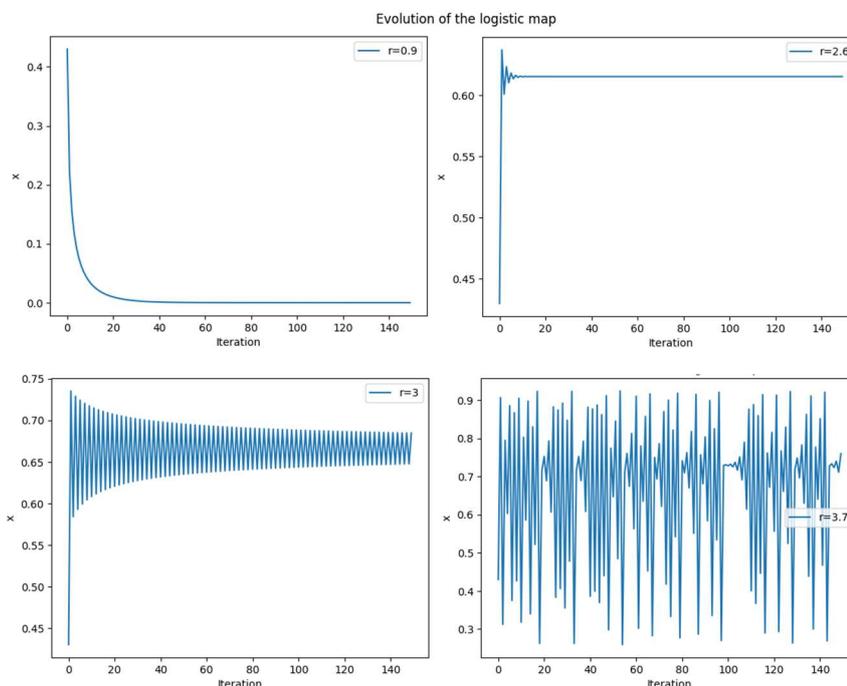


Gráfico 3: População variando em termos de diferentes valores para r . Quando $r < 1$ a população tende a extinção. Quando de $r=2$ até $r=3$ a população passa por oscilações, mas tende a valor específico. À medida que o valor de r se aproxima de 4, as oscilações se tornam caóticas. Código disponível no Apêndice A. Elaboração do Autor.

embora extremamente sensível às condições iniciais, exibe certos padrões e propriedades, como estruturas fractais, que podem ser caracterizados matematicamente. (MITCHELL, 2009)

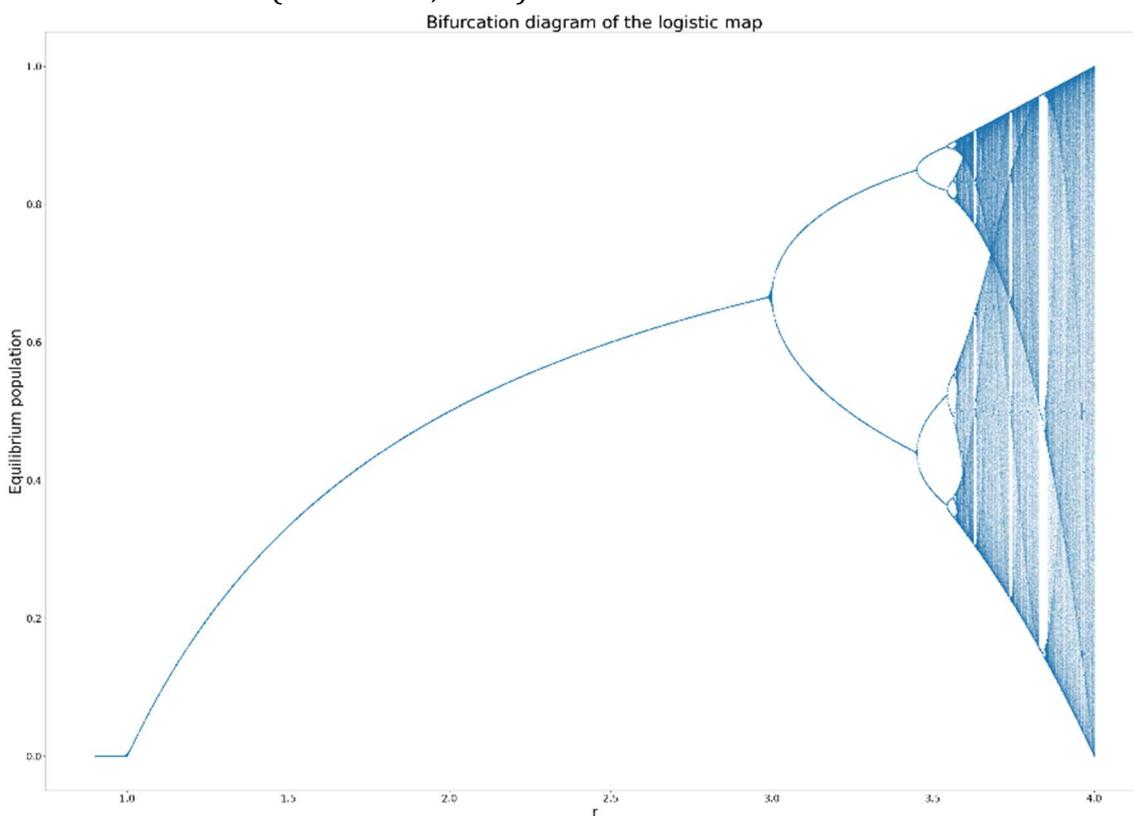


Gráfico 4: Mapa das bifurcações, para a mesma população do gráfico anterior. As oscilações são representadas como bifurcações que vão se abrindo a medida que o valor de r aumenta. Código disponível no Apêndice A. Elaboração do autor.

2.2 *Mr. Maxwell tem um demônio*

A entropia é um conceito fundamental da termodinâmica, estabelecido por Rudolf Clausius (1822 - 1888), como medida da qualidade da energia e se relaciona ao teorema de Carnot, que estabelece limites sobre a eficiência de máquinas térmicas (PINGUELLI, 2008). Na termodinâmica, a entropia representa a direção espontânea de processos e a predisposição de sistemas para atingir o equilíbrio. Na prática, os processos comumente encontrados são os irreversíveis: os sistemas evoluem sem a possibilidade de retornar aos seus estados iniciais. Nestas situações são considerados como sistemas isolados – sistemas que não trocam energia nem matéria com o ambiente – e, portanto, a entropia sempre aumenta. Este aumento na entropia representa uma degradação da energia, tornando-a menos acessível para uso em processos subsequentes de transformação, isto é, possuem menor capacidade de realizar trabalho (PINGUELLI, 2008).

Em um cenário ideal, durante um processo reversível – um processo que pode progredir em qualquer direção sem uma influência externa – a entropia de um sistema isolado permaneceria constante. No entanto, é essencial destacar que todos os processos do mundo real são irreversíveis em certa medida e envolvem invariavelmente um aumento na entropia.

A segunda lei da termodinâmica estipula que a entropia total de um sistema isolado tende a aumentar ou, no máximo, a permanecer constante ao longo do tempo. Este princípio pode se aplicar a sistemas abertos também ao se tomar como perspectiva que um sistema menor está sempre inserido em um ambiente e, o conjunto, pode ser visto como um sistema isolado – como se acredita que é o caso do nosso universo (PRIGOGINE, 1988). Portanto, dentro de um sistema fechado ou isolado, a dispersão total de energia, representada pela entropia, continua a crescer ou permanece constante. Esta lei é um princípio fundamental que rege a evolução dos sistemas termodinâmicos e destaca a direcionalidade inerente e irreversibilidade dos processos físicos com respeito à entropia, fornecendo uma das bases para a nossa compreensão da direção unidirecional do tempo, do passado para o futuro, conhecida como a *seta do tempo* (PRIGOGINE, 1988).

James Clerk Maxwell (1831 – 1879) propôs um experimento mental que ficou conhecido como o demônio de Maxwell, que desafiaria as implicações da segunda lei da termodinâmica. O demônio é concebido como um ser hipotético, inteligente, capaz de monitorar partículas individuais dentro de um sistema e permitir seletivamente apenas a passagem de partículas de alta energia através de uma partição, enquanto impede que partículas de baixa energia a atravessem. Ao fazer isso, o demônio poderia reduzir a entropia do sistema, contradizendo aparentemente a segunda lei (PINGUELLI, 2008).

Esse experimento mental levanta questões sobre a relação entre informação e entropia. O demônio de Maxwell parece possuir informações sobre as energias das partículas, permitindo-lhe manipular o sistema de forma a diminuir a entropia. No entanto, para operar, o demônio em si deve realizar trabalho e coletar informações, levando a um aumento de entropia em outros lugares, como em sua própria memória ou através da dissipação de energia. Em outras palavras, as ações seletivas do demônio exigem um gasto de energia e processamento de informações, levando, em última instância, a um aumento geral da entropia ao considerar o sistema que,

como um todo, inclui o demônio (PINGUELLI, 2008). Portanto, reafirmando a segunda lei da termodinâmica.

O experimento mental do demônio de Maxwell estimulou discussões sobre o papel da informação e da computação em sistemas físicos, introduzindo conceitos relacionados à teoria da informação e aos limites do conhecimento e da previsibilidade.

Ludwig Eduard Boltzmann (1844 – 1906) desenvolveu métodos estatísticos para compreender a distribuição das energias e posições das partículas, estabelecendo, assim, relações fundamentais entre entropia e a probabilidade de diferentes estados microscópicos. Neste contexto, a entropia fornece uma medida da dispersão de energia ou o número de configurações microscópicas (microestados) que correspondem ao estado macroscópico (macroestado) de um sistema. Essa correlação deriva do conceito de que, embora as trajetórias de partículas individuais possam variar de maneira determinística ou caótica, a abordagem coletiva de uma grande quantidade de partículas tende a revelar uma regularidade estatística (PINGUELLI, 2008).

A concepção teórica que Boltzmann desenvolveu viabilizou uma compreensão mais profunda das conexões entre os elementos observáveis em escala macroscópica e a dinâmica microscópica subjacente aos sistemas (PINGUELLI, 2008). Desta maneira, fenômenos macroscópicos como temperatura e pressão seriam, portanto, fenômenos emergentes a partir dessas interações microscópicas entre “átomos” de gás no sistema (referindo-se ao que hoje denominamos moléculas).

No contexto da época em que ele realizou suas pesquisas, final do século XIX e início do século XX, a existência dos átomos era uma hipótese controversa e mesmo desacreditada por alguns cientistas, principalmente Ernst Mach (1838 – 1916) e Wilhelm Ostwald (1853 – 1932), devido à sua natureza intrinsecamente invisível. O paradigma positivista predominante rejeitava afirmações sobre a existência de entidades inobserváveis, considerando-as como metafísicas e, portanto, fora do escopo da ciência (BRODA, 1981). No entanto, por meio de sua abordagem estatística e do estudo do comportamento dos gases, Boltzmann foi capaz de fornecer evidências persuasivas em apoio à existência dos átomos, demonstrando que a compreensão das propriedades macroscópicas da matéria poderia ser

explicada através do comportamento estatístico dos átomos individuais. Essa conquista foi fundamental para o estabelecimento do atomismo como um quadro científico legítimo, abrindo caminho para avanços adicionais na compreensão da natureza atômica da matéria que se seguiram e marcaram a física do século XX.

A mecânica estatística possibilitou o estudo de sistemas com um vasto número de graus de liberdade, fornecendo instrumental para o entendimento do comportamento de gases, do conceito de equilíbrio termodinâmico e da entropia. No entanto, como Melanie Mitchel (2009) reporta, há um problema com a abordagem estatística, pois ela fornece apenas o comportamento *provável* do sistema:

Por exemplo, se todas as moléculas de ar em um quarto estiverem se movendo aleatoriamente, é mais provável que elas estejam espalhadas por todo o ambiente, garantindo que todos tenham ar suficiente para respirar. Essa é a previsão em que confiamos e dependemos, e até hoje nunca nos decepcionou. No entanto, de acordo com a mecânica estatística, como as moléculas estão se movendo aleatoriamente, existe uma pequena chance de que, em algum momento, todas elas voem para o mesmo canto ao mesmo tempo. Nesse caso, qualquer pessoa que estivesse nesse canto seria esmagada pela enorme pressão do ar, e o restante de nós sufocaria por falta de ar. (MITCHELL, 2009, p.48-49. Tradução nossa.)

Um evento assim, porém, é extremamente improvável de acontecer, mesmo que, de acordo com a física newtoniana, exista uma probabilidade maior que zero para tal evento. A este respeito, Wiener comenta:

O próprio Gibbs acreditava que em um sistema no qual todos os invariantes haviam sido removidos como coordenadas extras, a maioria dos caminhos de pontos no espaço de fase passava por todas as coordenadas desse espaço. Essa hipótese ele chamou de hipótese ergódica, a partir das palavras gregas *ἔργον*, que significa "trabalho", e *ὁδός*, que significa "caminho". Agora, em primeiro lugar, como Plancherel e outros demonstraram, não há nenhum caso significativo em que essa hipótese seja verdadeira. Nenhum caminho diferenciável pode cobrir uma área no plano, mesmo que seja de comprimento infinito. Os seguidores de Gibbs, incluindo talvez Gibbs ele próprio no final, perceberam isso de forma vaga e substituíram essa hipótese pela hipótese quasi-ergódica, que simplesmente afirma que ao longo do tempo um sistema geralmente passa indefinidamente próximo a cada ponto na região do espaço de fase determinada pelos invariantes conhecidos. Não há dificuldade lógica quanto à verdade disso: é simplesmente totalmente inadequado para as conclusões que Gibbs baseia nele. Isso não diz nada sobre o tempo relativo que o sistema passa nas proximidades de cada ponto (WIENER, 1961, p.49.). Tradução nossa.

2.3 Ressonâncias rítmicas: Cibernética

As equações de Lotka-Volterra (Eq. 4) são duas equações diferenciais não lineares de primeira ordem que descrevem a dinâmica de sistemas biológicos nos quais duas espécies interagem, uma como predador e a outra como presa. As equações foram propostas independentemente por Alfred J. Lotka (1880 – 1949) em 1925 e Vito Volterra (1860 - 1940) em 1926 (HJORTH-JENSEN, 2009):

$$\frac{dx}{dt} = \alpha x - \beta xy$$

$$\frac{dy}{dt} = \delta xy - \gamma y$$

Eq. 4: Equações de Lotka-Volterra. Fonte: adaptado de (HJORTH-JENSEN, 2009)

Estas equações ilustram como, em sistemas ecológicos, as relações de predador-presa (Gráfico 5) também apresentam princípios cibernéticos. Isto é, quando a população de presas aumenta, há mais alimento para os predadores, e sua população também aumenta. À medida que a população de predadores cresce, eles consomem mais presas, reduzindo a população destas. Essa diminuição de alimento disponível, por sua vez, reduz a população de predadores. Esse ciclo continua, mantendo um equilíbrio dinâmico no ecossistema.

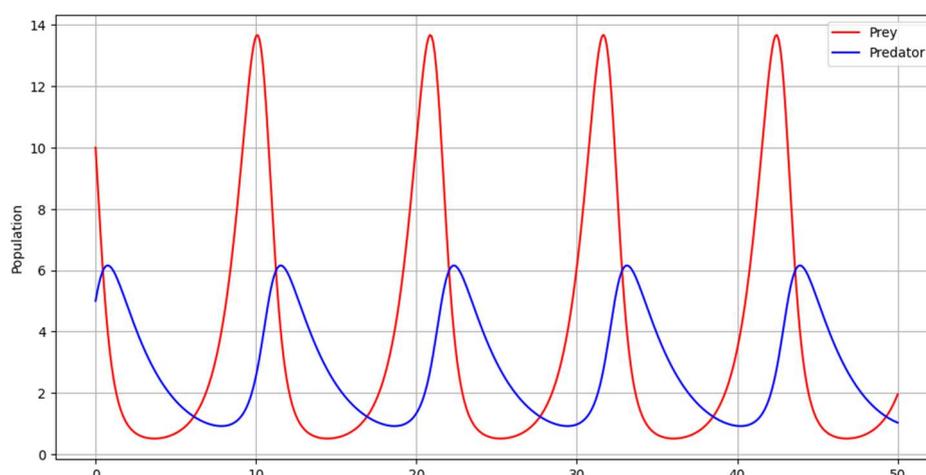


Gráfico 5: Interações decorrente do sistema Lotka-Volterra. Código disponível no Apêndice A. elaboração do autor.

Em um de seus textos de 1967, intitulado *A Growth Cycle*, o economista Richard M. Goodwin (1913 – 1996) utilizou as equações de Volterra para formalizar a teoria dos ciclos econômicos de Karl Marx (1818 – 1883). Ao aplicar as equações matemáticas fundamentadas na dinâmica populacional, Goodwin pretendia

fornecer uma base quantitativa para compreender a ciclicidade inerente observada nos sistemas econômicos.

Seu modelo voltou-se para as dinâmicas intrincadas entre diferentes setores econômicos e os mecanismos de retroalimentação que impulsionam as flutuações das variáveis-chave ao longo do tempo. A abordagem de Goodwin foi além dos modelos econômicos tradicionais – que dependiam de choques exógenos ou parâmetros estruturalmente instáveis –, ao demonstrar que o surgimento de ciclos autossustentáveis na atividade econômica podia ser endógeno, isto é, serem não dependentes de fatores externos para serem iniciados ou perpetuados (GOODWIN, 1982).

Essas ideias de processos de retroalimentação regulatória em sistemas é um dos conceitos fundamentais da cibernética. Em 1943, dois artigos seminais foram publicados: *Behavior, Purpose and Teleology* (ROSENBLUTH, WIENER, *et al.*, 1943), no qual se examina como os mecanismos de retroalimentação desempenham um papel crucial na manutenção da estabilidade, adaptabilidade e comportamento orientado por propósitos tanto em organismos biológicos quanto em máquinas. O outro artigo foi *A Logical Calculus of the Ideas Immanent in Nervous Activity* (MCCULLOCH, PITTS, 1943). Neste trabalho, os autores propõem um modelo de neurônio como um dispositivo lógico binário, introduzem o conceito de redes neurais compostas por esses neurônios, demonstram sua capacidade de computar qualquer função lógica e propõem sua aplicação à neurofisiologia e à psicologia.

Norbert Wiener (1894 - 1964) foi o responsável pela popularização do termo cibernética³, com a publicação de um livro de mesmo nome em 1948: *Cybernetics: or Control and Communication in the Animal and the Machine*. Este livro é uma coleção de textos independentes nos quais lança o problema de se tentar explicar processos complexos através das semelhanças entre os fenômenos biológicos e tecnológicos (WIENER, 1961).

Um dos principais temas da cibernética é o conceito de *sistemas de controle* que são regidos por um circuito de retroalimentação, ou *feedback*. O feedback, como um mecanismo de transferência de informação, assegura que um sistema mantenha

³ O termo, do grego antigo *κυβερνητικης*, significando *bom ao dirigir*, aparece na metáfora do timoneiro de Platão e é utilizada para simbolizar a arte política de governar. Maxwell fez uso deste termo para cunhar o termo *governador* como mecanismo de controle das máquinas à vapor, em um artigo da Royal Society de 1868 (KANG, 2016).

um estado específico. O processo envolve três etapas: a) o monitoramento contínuo do estado atual do sistema; b) a comparação com o estado desejado; e c) o ajuste subsequente com base na diferença percebida.

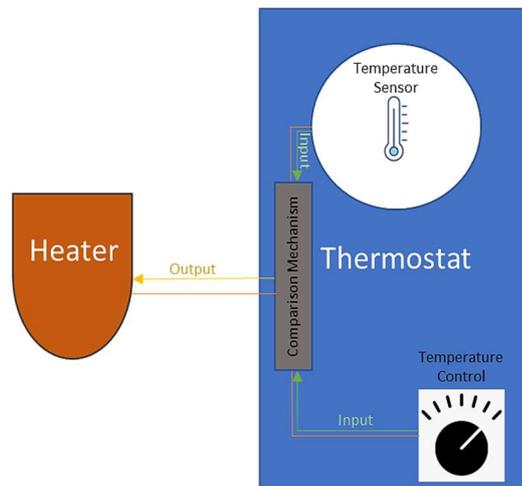


Figura 1: Representação esquemática de um termostato, ressaltando as setas de troca de informação (input) e controle (output). Fonte: elaboração do autor

Um dos exemplos mais frequentes de um sistema cibernético é um sistema de aquecimento controlado por um termostato. O termostato monitora continuamente a temperatura (informação de entrada), compara-a com a temperatura desejada pré-estabelecida e controla o sistema de aquecimento para diminuir ou aumentar o calor (*feedback*) de acordo com essa comparação. O objetivo é manter uma temperatura constante e confortável em um ambiente, independentemente das variações de temperatura externas.

Embora simples, este modelo incorpora características que poderiam descrever o comportamento de sistemas complexos. O corpo humano é um intrincado sistema cibernético. Por exemplo, considere a regulação dos níveis de açúcar no sangue no corpo. O pâncreas monitora os níveis de glicose no sangue. Quando esses níveis aumentam, o pâncreas secreta insulina, que estimula as células a absorverem glicose, reduzindo assim os níveis. Por outro lado, quando os níveis de glicose diminuem, o pâncreas libera glucagon, estimulando o fígado a liberar glicose armazenada. Esse ciclo de feedback mantém a homeostase, um estado interno constante. Esses laços retroalimentativos, ou *loops*, são críticos em um sistema onde as ações e decisões de uma parte influenciam reciprocamente e são influenciadas por outras partes do mesmo sistema.

A cibernética enfatiza, simultaneamente, o papel crucial da comunicação dentro de sistemas complexos. A natureza dinâmica da informação, caracterizada

por compartilhamento, troca e atualização contínuos em loops muitas vezes leva a comportamentos emergentes. Neste sentido, o estudo da cibernética intersecta os campos da complexidade, do caos e da teoria da informação, e um dos principais desafios para se entender o fluxo de informações é a questão do ruído.

Quando um sinal é transmitido, ele pode ser corrompido pelo ruído durante o processo de transmissão. O ruído introduz distorções, erros e perdas de informação no sinal original, o que pode afetar a qualidade e a precisão da mensagem recebida. E, sem risco de exagero, os sistemas biológicos são dominados pelo ruído, no entanto, seus sistemas são extremamente robustos.

2.4 Átomos de Informação

Os bits, unidades fundamentais de informação, pavimentam a era digital em que vivemos. Eles permeiam nossas vidas, desde os dispositivos eletrônicos que usamos até os sistemas mais elaborados que governam a sociedade. No entanto, além de sua onipresença intangível, bits são também portadores de significado, representando ideias, conhecimentos e, talvez – já houve quem o defendeu –, a própria essência do pensamento humano.

Apesar do progresso significativo na compreensão da complexidade e do caos, ainda existem muitas perguntas sem resposta. Uma das áreas importantes ainda a ser propriamente explorada está no campo da informação – como ela é armazenada, transferida e processada dentro de sistemas complexos. Acima de tudo, como ela é preservada através destes processos ao invés de se dissipar no ruído do meio.

Essencialmente, um sistema complexo é caracterizado por numerosos componentes interagindo, o que pode ser visto como cada um potencialmente carregando informações: podem ser o estado de uma partícula em um sistema físico, a saúde de uma célula em um sistema biológico, o sentimento de um indivíduo em um sistema social, assim por diante. Essas informações muitas vezes são interdependentes e mudam ao longo do tempo e à medida que interagem entre si.

No entanto, os mecanismos pelos quais as informações são armazenadas (como e onde os dados são retidos), transferidas (como os dados se movem entre os componentes) e processadas (como os dados são utilizados ou mudam de estado) dentro desses sistemas não são totalmente compreendidos. Essa falta de

compreensão decorre da própria complexidade desses sistemas. As relações não lineares, os laços de retroalimentação e as propriedades emergentes que caracterizam os sistemas complexos limitam os meios de rastrear o caminho das informações. Além disso, a sensibilidade desses sistemas às condições iniciais adiciona uma camada adicional de dificuldade na previsão de como as informações irão se desenvolver no decorrer do tempo.

O comportamento e a evolução dos sistemas complexos são amplamente impulsionados pelo fluxo e processamento de informações. Como um sistema responde a mudanças, se adapta ao longo do tempo e se organiza, muitas vezes depende de como as informações são tratadas dentro do sistema.

Norbert Wiener desenvolveu uma solução para o problema de estimativa de sinal e redução de ruído que funcionaria como um filtro. Seu trabalho foi motivado pelos esforços de guerra e permaneceu secreto até o final da Segunda Guerra Mundial. A formação matemática de Wiener e seu entendimento da teoria da probabilidade desempenharam um papel crucial no desenvolvimento do que ficou conhecido como filtro de Wiener. Ele formulou o filtro com base nos princípios de estimativa estatística e processamento de sinal, que funciona de maneira similar à minimização quadrada de erro, na verdade, podendo ser reduzido a esta. O trabalho inovador de Wiener estabeleceu as bases para várias aplicações em áreas como telecomunicações, processamento de imagem e aprimoramento de áudio.

Walter Pitts (1923 – 1969) e Warren McCulloch (1898 – 1969) procuravam uma solução digital para o problema da troca e processamento de informações.

McCulloch, na época em que conheceu Pitts, tinha 42 anos e era um filósofo-poeta confiante, de olhos cinzentos, barba selvagem e fumante inveterado, que vivia de uísque e sorvete e nunca ia dormir antes das 4 da manhã. Já Pitts, então com 18 anos, era pequeno e tímido, com uma testa comprida que o fazia parecer mais velho e um rosto rechonchudo, parecido com um pato de óculos. McCulloch era um cientista respeitado. Pitts era um jovem fugido de casa e sem-teto. Ele havia estado por perto da Universidade de Chicago, trabalhando em um emprego insignificante e se infiltrando nas palestras de Russell. Foi na universidade que conheceu um jovem estudante de medicina chamado Jerome Lettvin. Lettvin foi quem apresentou os dois homens. (...) Embora tenham começado em extremos opostos do espectro socioeconômico, McCulloch e Pitts estavam destinados a viver, trabalhar e morrer juntos.(GEFTER, 2015. sp. Tradução nossa.)

Ambos haviam sido inspirados pelo *Principia*⁴, e imaginavam que o sistema nervoso deveria obedecer a uma estrutura lógica construída a partir de princípios simples. Neurônios se comportariam assim como operadores lógicos operando sobre proposições (Gráfico 6), a aritmética modular permitia se estabelecer loops entre os estados. A máquina de Turing poderia, desta maneira, servir como o modelo para se compreender a mente humana (MCCULLOCH, PITTS, 1943). Wiener propôs a Pitts como projeto de pesquisa de doutoramento que tornasse esta hipótese mais realista, possibilitando eventualmente produzir uma rede neural artificial que simulasse o comportamento observado no cérebro humano. (GEFTER, 2015)

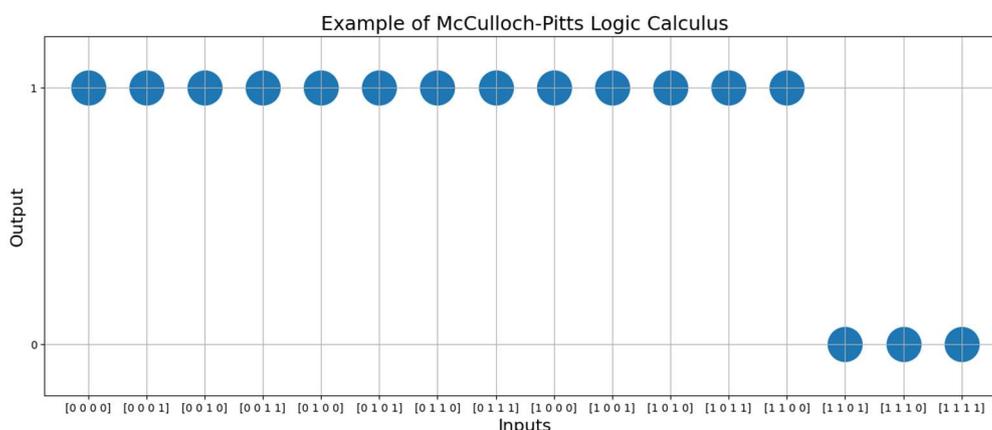


Gráfico 6: Representação de inputs e outputs para uma rede neural de três camadas, uma oculta. Código disponível no Apêndice A. Elaboração do autor.

Embora, na época, a estrutura do DNA não houvesse sido totalmente elucidada, já se tinha conhecimento da quantidade de informações que o DNA poderia codificar. Porém, ao considerar o elevado número de conexões neuronais, da ordem de trilhões no cérebro humano, tornava-se evidente que era impraticável armazenar diretamente essa imensa quantidade de informação no código genético. Portanto, era plausível que algum mecanismo de compressão de informações estivesse em ação.

Conseqüentemente, Pitts propôs que os cérebros inicialmente são formados com redes neurais aleatórias, que possuem estados altamente prováveis e mínima quantidade de informação (uma noção que ainda hoje é discutida). Ele supunha que,

⁴ *Principia Mathematica*, escrito por Bertrand Russell em parceria com Alfred Whitehead, o livro foi publicado entre 1910 e 1913 e busca estabelecer uma base lógica sólida para os fundamentos da matemática. Os autores propuseram demonstrar que todos os princípios matemáticos podem ser derivados de proposições lógicas básicas. Eles usaram a lógica simbólica para construir uma estrutura rigorosa e formal que permitisse a dedução matemática a partir de princípios lógicos fundamentais. O objetivo era fornecer uma fundamentação sólida e consistente para toda a matemática, evitando contradições e paradoxos (SIGMUND, HOFSTADTER, 2017).

ao ajustar os limiares de ação dos neurônios ao longo do tempo, as redes neurais aleatórias realizavam a transição do caos para a ordem, em outras palavras, haveria um mecanismo que permitiria que a informação emergisse através da auto organização (GEFTER, 2015). Esse conceito sugere, portanto, que o cérebro, ao modificar os limiares de ativação dos neurônios, pode se organizar com base em estímulos externos e padrões de experiências, adquirindo a capacidade de aprender e de dar sentido aos seus conteúdos.

As ideias de Pitts e McCulloch não receberam muita atenção dos neurocientistas da época, mas eram por demais irresistíveis no meio dos pioneiros da computação. Em 1945, John Von Neumann (1903 – 1957) sugeriu o uso de abstração para a programação de computadores (NEUMANN, 1993). Naquela época, a programação era realizada manualmente através do cabeamento dos circuitos. No entanto, Von Neumann propôs uma mudança significativa: utilizar válvulas como interruptores lógicos para substituir o cabeamento manual. Essas válvulas realizariam operações similares aos cálculos realizados pelos neurônios de Pitts e McCulloch, estabelecendo uma conexão entre o funcionamento dos computadores e o processamento neural (GEFTER, 2015). A memória seria realizada por conexões em loops de aritmética modular, sendo um tipo de elemento que estimula a si mesmo e capaz de sustentar este estímulo indefinidamente (NEUMANN, 1993). Essa abstração permitiria uma programação mais eficiente e flexível, abrindo caminho para o desenvolvimento dos computadores modernos.

Em 1957, Frank Rosenblatt (1928 – 1971), psicólogo e cientista da computação americano, introduziu o *Perceptron*. Inspirado no entendimento dos sistemas neurais biológicos na época, Rosenblatt tinha como objetivo criar uma máquina capaz de imitar certos aspectos do aprendizado biológico (ROSENBLATT, 1958). Perceptrons com uma camada de neurônios (*single layer*) eram capazes de realizar classificação de padrões linearmente separáveis e a adição de novas camadas (*multi-layer*), supunha-se então, aumentaria a capacidade de aprender. Porém isto estava tecnologicamente além do que era possível implementar em sua época e apenas realizado propriamente após a introdução de técnicas de *deep learning* (LECUN, BENGIO, *et al.*, 2015).

Pitts, em sua tese doutoral sobre redes neurais probabilísticas tridimensionais, tinha como objetivo utilizar os princípios da mecânica estatística

para modelar a natureza probabilística da atividade neural, revelando os mecanismos pelos quais a informação poderia emergir produzindo processos cognitivos complexos.

Por meio da mecânica estatística, ele pretendia descrever a evolução dos sistemas em direção a estados de maior organização e menor energia, explicando como ocorre a reorganização gradual de conexões inicialmente aleatórias entre os neurônios em padrões mais estruturados. Esse modelo explicaria como interações locais entre neurônios dão origem a propriedades globais emergentes nas redes neurais (DYSON, 2012).

Ao se apropriar de conceitos como o de entropia, a mecânica estatística pode avaliar a complexidade, eficiência e a capacidade dos sistemas neurais. Essas ferramentas permitem uma compreensão mais profunda de como a informação é codificada, transmitida e manipulada dentro da rede, revelando os processos fundamentais subjacentes à cognição que Pitts procurava desvendar.

Além disso, a mecânica estatística facilitaria a exploração de transições de fase e fenômenos críticos dentro de seu modelo de cérebro lógico-digital. Transições de fase representam *pontos críticos* nos quais o sistema passa por mudanças significativas no comportamento ou na organização. Essas transições podem representar, no cérebro humano, por exemplo, períodos críticos durante o desenvolvimento cerebral cuja falha poderia estar associada ao aparecimento de deficiências cognitivas ou doenças mentais; ou a emergência de funções cognitivas específicas, como o período de aquisição da linguagem. Por meio do uso de conceitos como cadeias de Markov, processos estocásticos e distribuições de equilíbrio, a mecânica estatística também possibilitaria a modelagem de dinâmicas temporais na atividade neural, proporcionando uma compreensão mais profunda dos processos de aprendizagem, memória e processamento de informações.

Durante o período de escrita de sua tese, Pitts colaborou com o biólogo Humberto Maturana, além de McCulloch, envolvendo uma série de experimentos com rãs, cujo objetivo era compreender como o sistema visual desses animais processa e interpreta informações (LETTVIN, MATURANA, *et al.*, 1959).

Os pesquisadores expuseram os animais a uma variedade de experiências visuais, manipulando diversos aspectos do ambiente. Eles experimentaram aumentar e diminuir a intensidade da luz, mostraram às rãs fotografias coloridas

retratando seu habitat natural e utilizaram ímãs para suspender moscas artificiais, criando movimentos simulados de presas (GEFTER, 2015, LETTVIN, MATURANA, *et al.*, 1959).

Durante esses experimentos, a equipe concentrou-se em medir as respostas dos olhos rãs antes que as informações fossem transmitidas ao cérebro.

Mas o que eles descobriram os surpreendeu: contrariando a crença predominante de que o olho atuava apenas como um receptor passivo de estímulos visuais, digamos, como uma lente de uma câmera, os pesquisadores constataram que o olho da rã estava ativamente envolvido na filtragem e análise de informações visuais. Em outras palavras, os olhos não apenas registram o que veem, mas extraem características visuais específicas, como contraste, curvatura e movimento (GEFTER, 2015).

Em seu artigo inovador intitulado *What the Frog's Eye Tells the Frog's Brain*, publicado em 1959, Pitts, McCulloch e Maturana detalharam suas descobertas. A pesquisa revelou que o olho da rã se comunicava com o cérebro usando uma forma que já estava altamente organizada e interpretada. Não se tratava de uma simples transmissão de dados visuais brutos, o olho desempenhava um papel ativo no processamento e na extração de informações significativas antes de enviá-las ao cérebro para processamento adicional (LETTVIN, MATURANA, *et al.*, 1959).

Estas descobertas foram exatamente o oposto do que Pitts esperava encontrar, ao procurar um cérebro que receberia informações brutas e procederia através de passos lógicos e hierárquicos processando esta informação, Pitts se deparou com um ruído caótico onde o processamento era distribuído e que, de alguma maneira, funcionava. Neste momento sua tese foi refutada⁵, a lógica que considerava o bloco fundamental do sistema nervoso não parecia ser muito relevante, isso, se de fato existisse. O cérebro não poderia ser, portanto, redutível uma máquina de Turing.

⁵ Em vista disto, Pitts nunca concluiu o doutorado – na verdade, nunca concluiu os estudos de nível médio. Faleceu em 1969 (dez anos depois da publicação do artigo em pauta), aos 46 anos de idade, em decorrência do uso excessivo de álcool, enquanto morava em um quarto de pensão. McCulloch faleceu poucos meses depois, aos 70 anos de idade, por problemas cardíacos (GEFTER, 2015).

3 TEORIA DA INFORMAÇÃO

Todos começamos a partir do "realismo ingênuo", isto é, a doutrina de que as coisas são como parecem. Pensamos que a grama é verde, que as pedras são duras e que a neve é fria. Mas a física nos assegura que o verde da grama, a dureza das pedras e a frieza da neve não são o verde da grama, a dureza das pedras e a frieza da neve que conhecemos em nossa própria experiência, mas algo muito diferente. No que segue, vamos observar a vida através da lente da aleatoriedade e ver que muitos dos eventos de nossas vidas, também, não são exatamente o que parecem, mas sim algo muito diferente.

Bertrand Russell

Nesta seção, uma análise das contribuições relevantes de três proeminentes pesquisadores será conduzida, com um destaque específico sobre como seus conceitos foram instrumentalmente decisivos na formação da compreensão contemporânea de informação, computação e sistemas complexos.

Andrey Markov, através da introdução das cadeias de Markov, forneceu instrumentos analíticos fundamentais para a modelagem e a compreensão de sequências e suas propriedades probabilísticas. A propriedade de Markov, que postula que o estado futuro é exclusivamente dependente do estado atual, tem sido empregada em uma vasta gama de estudos de sistemas complexos, desde a análise de padrões de letras em textos literários até a investigação de comportamentos emergentes em sistemas biológicos. Claude Shannon, pioneiro na introdução da teoria da informação, inaugurou um novo campo acadêmico que busca compreender os fundamentos da transmissão de informações. A equação de entropia de Shannon, um instrumento para quantificar a incerteza inerente à quantidade de informação contida em uma mensagem, juntamente com a aplicação do conceito de entropia para representar a aleatoriedade da informação, configuram a base estrutural para a construção da comunicação digital contemporânea. Alan Turing, através da proposição da Máquina de Turing e do enunciado do Problema da Parada, estabeleceu um marco teórico para a análise dos algoritmos e para explorar as potencialidades (bem como as limitações) intrínsecas ao domínio da computação. Ao amalgamar todas essas ideias, este capítulo propõe uma exploração abrangente das dinâmicas subjacentes à teoria da complexidade.

3.1 O caminho menos percorrido

As cadeias de Markov receberam o nome do matemático russo Andrey Markov (1856 – 1922), que introduziu o conceito no início do século XX. Markov estava interessado em compreender e modelar as probabilidades de eventos em sequências que pareciam ter um certo grau de aleatoriedade, mas ainda exibiam padrões e dependências. O trabalho inicial de Markov concentrou-se na análise de sequências de letras na língua russa e no estudo das probabilidades de diferentes letras aparecerem uma após a outra (BASHARIN, LANGVILLE, *et al.*, 2004). Ele desenvolveu modelos matemáticos conhecidos como cadeias de Markov para descrever essas sequências e suas propriedades probabilísticas:

No que se tornou a famosa primeira aplicação das cadeias de Markov, A. A. Markov estudou a sequência de 20.000 letras no poema *Eugeny Onegin* de A. S. Pushkin, descobrindo que a probabilidade estacionária de vogais é $p = 0,432$, que a probabilidade de uma vogal após uma vogal é $p_1 = 0,128$, e que a probabilidade de uma vogal após uma consoante é $p_2 = 0,663$. No mesmo artigo, Markov também apresentou os resultados de seus outros testes; ele estudou a sequência de 100.000 letras no romance *The Childhood of Bagrov, the Grandson* de S. T. Aksakov. Para esse romance, as probabilidades foram $p = 0,449$, $p_1 = 0,552$ e $p_2 = 0,365$. (BASHARIN, LANGVILLE, *et al.*, 2004. p. 1). Tradução nossa.

O estudo das cadeias de Markov ganhou maior destaque no meio do século XX, principalmente nos campos da matemática, estatística e ciência da computação. Os pesquisadores perceberam as amplas aplicações das cadeias de Markov na modelagem de diversos sistemas, desde processos físicos até comportamentos humanos. As cadeias de Markov são uma ferramenta para capturar essas dinâmicas temporais e transições probabilísticas entre diferentes estados. Ao representar sistemas complexos como uma série de estados e estudar as probabilidades de transição entre esses estados, as cadeias de Markov podem ajudar a elucidar comportamentos emergentes, padrões e a complexidade geral do sistema.

A ideia por detrás das cadeias de Markov é que o estado futuro depende apenas do estado atual e é independente dos estados anteriores, uma propriedade conhecida como propriedade de Markov. Essa propriedade simplifica a análise de sistemas complexos, concentrando-se no estado atual e em suas transições probabilísticas. Isso nos permite estudar o comportamento do sistema ao longo do tempo e examinar como diferentes estados influenciam os resultados futuros.

Vamos considerar um exemplo em que temos um estado atual chamado A e, no próximo instante, temos três opções: permanecer no estado A, mover para o estado B ou mover para o estado C. Para cada transição, podemos atribuir uma probabilidade específica. Por exemplo, podemos ter uma probabilidade de 20% de permanecer em A, 50% de ir para B e 30% de ir para C.

Essas probabilidades podem ser determinadas observando eventos anteriores. Por exemplo, ao analisar uma sequência de letras, podemos calcular a probabilidade de uma determinada letra ser seguida por outra letra específica com base nas ocorrências anteriores. Essa informação nos ajuda a entender e modelar as transições entre os estados, que poderíamos representar (Figura 2).

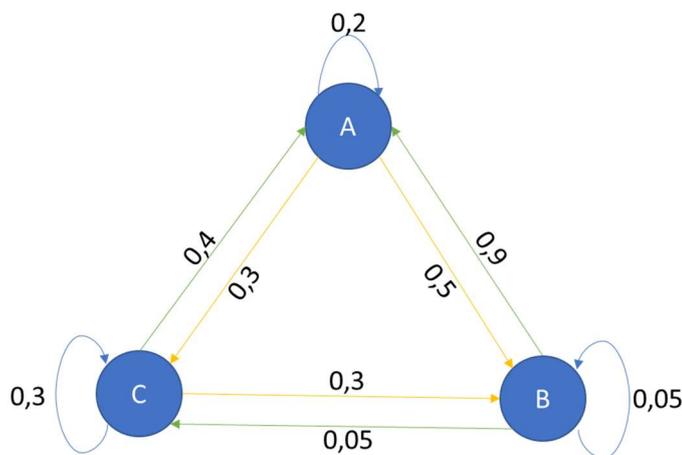


Figura 2: Transições de estados e probabilidades associadas. Fonte: elaboração do autor.

Uma abordagem para lidar com esses dados é representá-los como uma matriz quadrada. Cada elemento a_{ij} da matriz representa uma probabilidade de transição entre dois estados. Por exemplo, o elemento a_{11} representa a probabilidade de, estando no estado A, permanecer em A no próximo passo. Da mesma forma, o elemento a_{23} representa a probabilidade de, estando no estado B, ir para o estado C. Assim, podemos organizar essas probabilidades em uma matriz para representar as transições entre os estados:

$$M = \begin{bmatrix} 0,2 & 0,5 & 0,3 \\ 0,9 & 0,05 & 0,05 \\ 0,4 & 0,3 & 0,3 \end{bmatrix}_{3 \times 3}$$

Para encontrar as probabilidades estacionárias em um processo de Markov, podemos usar o conceito de *autovetores* e *autovalores*. Assim, podemos fazer $Mv = \lambda v$. No caso específico em que o autovalor $\lambda=1$, podemos calcular o autovetor

associado a esse autovalor para obter as probabilidades de longo prazo para cada estado. Essencialmente, a ideia é que, quando multiplicamos repetidamente a matriz de transição do processo de Markov por um vetor inicial (que pode ser um vetor de probabilidades aleatórias), eventualmente chegaremos a um estado estacionário $v = [-0,77 \quad -0,52 \quad -0,37]$. Esse estado estacionário é representado pelo autovetor associado ao autovalor 1. Normalizando o autovetor, obtemos as probabilidades de longo prazo para cada estado $v = [0,46 \quad 0,31 \quad 0,22]$.

Os valores mencionados podem ser obtidos por meio da iteração de um grande número de transições de estado. Para realizar essa tarefa, implementamos uma simulação em *Python*⁶ e permitimos que ela seja executada até que fique evidente que os valores das probabilidades se estabilizaram (Gráfico 7). Os valores obtidos foram computacionalmente⁷ foram: $v = [0,4640355 \quad 0,313898 \quad 0,2220665]$, que coincidem com os resultados anteriores.

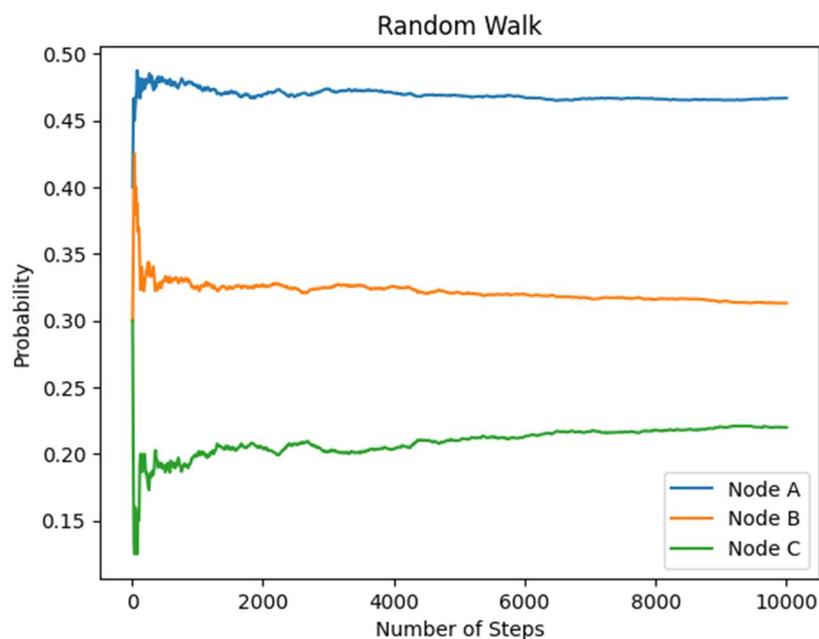


Gráfico 7: Randon Walk de 10.000 iterações. Código disponível no Apêndice A. elaboração do autor

O uso de cadeias de Markov no contexto da complexidade vai além de modelos probabilísticos simples. Elas têm sido aplicadas em diversas áreas, como biologia, economia, ciências sociais e ciência da computação, para compreender a dinâmica de ecossistemas, mercados financeiros e redes sociais. As cadeias de Markov

⁶ Código disponível no Apêndice A.

⁷ Através de *linalg.eig* da biblioteca Numpy, após normalização.

permitem explorar a interação complexa de múltiplas variáveis, dependências e retroalimentações dentro de um sistema sob um viés quantitativo, evidenciando como comportamentos intrincados que emergem a partir de regras probabilísticas simples

3.2 *Surpresas na transmissão*

Sob segredo, outra teoria estava sendo gestada durante a Segunda Guerra Mundial, que justamente se direcionava para o problema do ruído. Claude Shannon (1916 – 2001) escreveu um artigo, publicado em 1948, que incorporava a mecânica estatística no estudo da comunicação (SHANNON, 1948). Seu trabalho se concentrou na quantificação da informação e na determinação da taxa máxima na qual ela pode ser transmitida por um canal, como um fio ou qualquer outro meio. A definição de informação que utilizou baseia-se no conceito de entropia, que aqui mede a incerteza ou aleatoriedade de uma mensagem. De acordo com Shannon, a informação é uma redução na incerteza ou a resolução da incerteza. Nas palavras de Gregory Bateson (1904 – 1980) informação neste contexto seria *qualquer diferença que faça diferença* (BATESON, 1979).

Shannon utilizou-se das equações de Boltzmann, mas, ao invés de aplicá-las a gases, tentou entender como a ordem presente em uma informação se relaciona à aleatoriedade de todas as possíveis formas que os conteúdos que compõem esta mensagem poderiam estar organizados. De certa maneira, a mensagem seria entendida como uma propriedade emergente do sistema. Wiener comenta:

A noção de quantidade de informação vincula-se muito naturalmente a uma noção clássica da mecânica estatística: a de entropia. Assim como a quantidade de informação em um sistema é uma medida do seu grau de organização, então a entropia de um sistema é uma medida do seu grau de desorganização; e um é simplesmente o negativo do outro. Este ponto de vista nos leva a uma série de considerações sobre a segunda lei da termodinâmica, e para um estudo da possibilidade dos chamados demônios de Maxwell. Tais perguntas surgem independentemente no estudo de enzimas e outros catalisadores, e seu estudo é essencial para a compreensão adequada de tais fenômenos da matéria viva como metabolismo e reprodução (WIENER, 1961, p. 11. Tradução nossa.).

Seu modelo foi a transmissão das letras de uma mensagem através do telégrafo e teletipo e, portanto, trata-se de canais discretos, neste caso, símbolos.

Um canal discreto refere-se a um sistema em que uma sequência de escolhas de um conjunto finito de símbolos elementares pode ser transmitida de um ponto a

outro. Cada símbolo possui uma certa duração no tempo, e não é necessário que todas as sequências de símbolos possíveis sejam permitidas para transmissão no sistema. Por exemplo, no telégrafo, os símbolos podem ser definidos como:

- (1) um ponto representado por fechamento da linha por uma unidade de tempo seguida de abertura da linha por uma unidade de tempo;
- (2) um traço representado por três unidades de tempo de fechamento e uma unidade de abertura;
- (3) um espaço entre letras representado por três unidades de abertura da linha; e
- (4) um espaço entre palavras representado por seis unidades de abertura da linha. Certas restrições podem ser impostas às sequências de símbolos permitidas, como evitar espaços consecutivos para prevenir ambiguidade com um espaço entre palavras.(SHANNON, 1948. p.3. Tradução nossa.)

A questão que Shannon se coloca é como seria possível medir a capacidade de um canal discreto desse tipo. Capacidade refere-se à taxa máxima na qual informações podem ser transmitidas de forma confiável pelo canal. No caso de canais discretos, a capacidade é tipicamente medida em termos do número de símbolos que podem ser transmitidos por unidade de tempo. A medição da capacidade do canal envolve avaliar fatores como a duração de cada símbolo, as restrições nas sequências de símbolos e a confiabilidade da transmissão.

Ademais, ao se transmitir um texto em português, sabemos que nem todos os símbolos têm a mesma probabilidade de aparecer: alguns símbolos são mais comuns do que outros. Por exemplo, a letra "a" é geralmente mais utilizada (com uma frequência aproximada de 14,634%) em comparação com a letra "x" (com 0,253%). Portanto, ao transmitir um texto em português, é mais provável encontrar o símbolo "a" do que o símbolo "x" nesta mensagem. Além disso, certas combinações de símbolos ocorrem com mais frequência do que outras. Por exemplo, ao analisar um texto em português, é comum encontrar a combinação "qu" em palavras, como em "queijo" ou "quarto". Porém, combinações como "qt" ou "qv" são extremamente raras, devido às restrições da língua portuguesa. Ou seja, a ocorrência de certas combinações de símbolos pode não adicionar informações novas.

No seu artigo, Shannon propõe um método para gerar palavras e sentenças de forma probabilística levando em consideração a frequência de ocorrência de cada letra no idioma inglês. Ele utiliza uma abordagem chamada de *aproximação de ordem zero*, na qual cada símbolo tem a mesma probabilidade de ocorrer e não há dependência entre símbolos consecutivos (as probabilidades são independentes).

Para uma maior precisão, a *aproximação de primeira ordem*, incorpora as probabilidades calculadas para cada símbolo com base na língua inglesa, independentemente. Além disso, ele considerou as frequências de ocorrência de pares de letras (digramas) e grupos de três letras consecutivas (trigramas) na aproximação de segunda e terceira ordem, respectivamente. À medida que ordens mais elevadas são utilizadas para gerar aleatoriamente o texto, ele se torna cada vez mais semelhante ao inglês em termos de estrutura e padrões linguísticos (Quadro 2). A mesma abordagem foi utilizada na construção de sentenças, usando agora palavras como elementos básicos, como resultado a semelhança com frases do inglês aumenta significativamente em aproximações de ordem mais alta.

Quadro 2: Séries de aproximações ao inglês com letras extraídos de Shannon(1948), p.7.

<p>1. Zero-order approximation (symbols independent and equiprobable). XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZL-HJQD.</p> <p>2. First-order approximation (symbols independent but with frequencies of English text). OCRO HLI RGWR NMIELWIS EU LL NBNESBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL.</p> <p>3. Second-order approximation (digram structure as in English). ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TU-COOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE.</p> <p>4. Third-order approximation (trigram structure as in English). IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONS-TURES OF THE REPTAGIN IS REGOACTIONA OF CRE.</p>
--

Em vista disso, consideramos uma série de eventos possíveis, p_1, p_2, \dots, p_n , como a escolha de letras acima, desejamos quantificar a incerteza associada $H(p_1, p_2, \dots, p_n)$ que represente a quantidade de informação contida na mensagem. Para tanto, Shannon desenvolve uma função tal que exhibe:

1) Continuidade: H deve exibir continuidade em relação às probabilidades p_i . Isso significa que pequenas alterações nas probabilidades devem resultar em pequenas alterações no valor de H .

2) Monotonicidade: se todas as probabilidades p_i forem iguais (denotadas como $p_i = 1/n$), então H deve ser uma função monotonicamente crescente de n . Em termos mais simples, à medida que o número de eventos igualmente prováveis aumenta, deve haver mais escolha ou incerteza, resultando em um valor mais alto para H .

3) Aditividade: Se uma escolha for dividida em duas escolhas sucessivas, o valor original de H deve ser igual à soma ponderada dos valores individuais de H correspondentes a cada escolha. Em outras palavras, a medida de incerteza para a escolha original deve ser a medida combinada de incerteza para as duas escolhas sucessivas. Desta maneira ele chega à equação que ficou conhecida como Equação da Entropia de Shannon (Eq. 5), a base do logaritmo dois é em relação ao caso de uma escolha binária.

$$H = -\sum_{i=1}^n p_i \log_2 p_i$$

Eq. 5: Equação da Entropia de Shannon. Fonte: adaptado de Shannon (1948)

Ao observar o Gráfico 8, podemos notar que diferentes probabilidades se mapeiam de forma simétrica em relação a $p=0,5$, revelando distintos níveis de Entropia de Shannon. Na teoria da informação, quando as probabilidades dos eventos são distribuídas de maneira equitativa, significa que cada evento tem a mesma chance de ocorrer, resultando em um alto grau de incerteza ou entropia. Por exemplo, quando $p = 0,5$, as probabilidades de dois eventos, como "cara" e "coroa" em um lançamento de uma moeda honesta, são iguais. Nesse caso, temos uma incerteza máxima, pois não é possível prever com certeza o resultado. À medida que as probabilidades se afastam de $p = 0,5$, a incerteza diminui e o valor da entropia também diminui.

Por outro lado, quando $p = 0$ ou $p = 1$, o resultado se torna determinístico. Em outras palavras, se um evento é quase certo de acontecer, ele não fornece muitas informações novas ou esclarecedoras, pois sua ocorrência é esperada.

O conteúdo de informação de um evento está diretamente ligado ao nível de surpresa ou inesperado que ele traz. Quanto menor a probabilidade de ocorrência de um evento, maior será o seu conteúdo de informação. Eventos altamente improváveis possuem um conteúdo de informação mais elevado e à medida que a probabilidade de ocorrência do evento aumenta, o seu conteúdo de informação diminui gradualmente. Este resultado corrobora a lei de Zipf, proposta por George Kingsley Zipf (1902 – 1950) em 1935, que descreve a distribuição de frequência das palavras em uma linguagem natural, afirmando que um pequeno número de palavras ocorre com muita frequência, enquanto a maioria das palavras ocorre raramente. Desta maneira, as palavras mais comuns são altamente previsíveis e,

portanto, carregam menos informação; enquanto as palavras mais raras, em contraste, são menos previsíveis e, conseqüentemente, carregam mais informação.

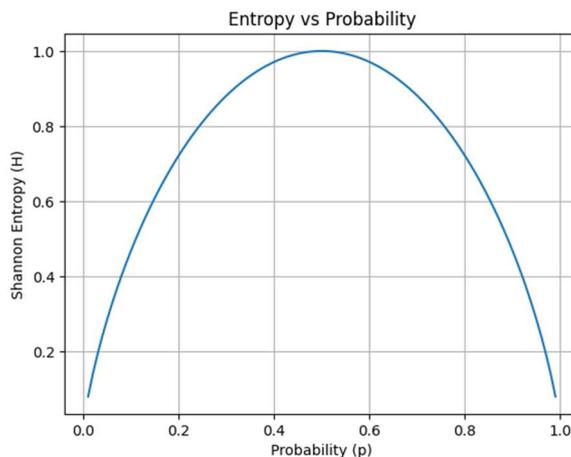


Gráfico 8: Relação entre entropia e informação, em uma escolha binária a informação é máxima quando a probabilidade para cada escolha é de $1/2$. Fonte: elaboração do autor.

Voltando ao nosso exemplo, se encontrarmos a combinação de letras "qt" em um texto, podemos inferir que ela está relacionada a uma fórmula ou a um nome próprio. Nesse caso, a omissão da letra "t" após o "q" seria mais prejudicial para a compreensão do texto do que a omissão da letra "u" após o "q".

Por fim, desejamos quantificar o que cada variável adiciona ao total de informação ganho. A fórmula nos mostra que em uma mensagem, por exemplo, com n letras, computamos a probabilidade de cada letra p_i e desejamos descobrir o total de conteúdo informativo na mensagem como um todo $H(p_1, p_2, \dots, p_n)$, conforme ilustrado no Gráfico 9.

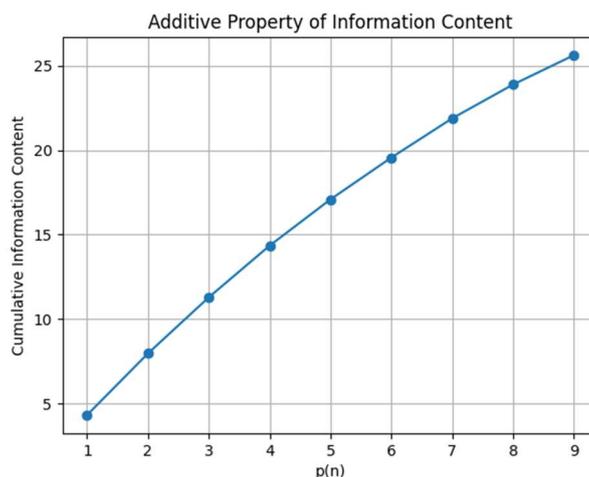


Gráfico 9: O gráfico ilustra a propriedade aditiva do conteúdo de informação. O conteúdo de informação aumenta à medida que o mais dados são recebidos, cada um acrescentando inversamente à probabilidade de sua ocorrência, onde $p_1 < p_2 < \dots < p_n$. Fonte: elaboração do autor.

3.3 Máquinas Inteligentes: Turing

A máquina de Turing (TM) é um modelo teórico de computação e cálculo mecânico introduzido pelo matemático britânico Alan Turing (1912 – 1954) em 1936. Turing propôs esse modelo como um experimento mental sobre os limites da computação. Embora não seja uma máquina física, é um modelo matemático que captura os elementos essenciais da computação (TURING, 1936).

Uma TM pode ser pensada como uma fita de memória infinita dividida em células, cada uma capaz de armazenar um símbolo. Essa fita é manipulada por um dispositivo ou *cabeçote* que pode ler e escrever símbolos na fita e mover a fita para a esquerda ou para a direita, uma célula de cada vez⁸ (Tabela 1).

Tabela 1: Operações de uma TM. Fonte: elaboração do autor.

Componente	Descrição
Fita	A fita é de comprimento infinito e dividida em células. Cada célula pode conter um símbolo ou ficar em branco.
Cabeçote de Leitura e Escrita	O cabeçote move-se ao longo da fita e pode ler o símbolo na célula em que se encontra. Ele também pode escrever um símbolo na célula ou apagar o símbolo nessa célula.
Registro de Estado	O registro de estado armazena o estado da máquina de Turing, que determina a operação a ser realizada (como leitura, escrita, movimento do cabeçote ou interrupção da operação).
Tabela de Instruções	Esta é essencialmente o 'programa' da máquina de Turing. Ela diz à máquina o que fazer com base no estado atual e no símbolo sendo lido. As instruções incluem escrever um símbolo, mover o cabeçote (para a esquerda ou para a direita) e fazer a transição para um novo estado.

Suponha que desejamos uma TM que incremente um número binário em 1. Nossa máquina precisa lidar com números binários, portanto, ela usará dois símbolos, '0' e '1', além de um símbolo em branco para representar o final do conteúdo de entrada (Tabela 2).

⁸ Um exemplo da implementação do algoritmo *busy beaver* pode ser encontrado em (KOVALSKI, 2021).

Tabela 2: Tabela de instruções para adição binária, equivalente ao que hoje chamaríamos de um programa. Fonte: elaboração do autor.

Current State	Symbol Read	Symbol to Write	Move Direction	Next State
Start	1	0	Right	Carry
Start	0	1	N/A	Halt
Start	Blank	1	N/A	Halt
Carry	1	0	Right	Carry
Carry	0	1	N/A	Halt
Carry	Blank	1	N/A	Halt

Tomando o número binário 1011, equivalente ao decimal 11, por exemplo⁹, o escrevemos na fita da direita para a esquerda, isto é, em ordem invertida, de tal forma que o dígito menos significativo está na primeira posição da fita. Então, o cabeçote inicia a leitura da fita, e, seguindo a descrição da tabela (Tabela 3), realiza, passo a passo, as operações até atingir o estado de parada (*halt state*). Neste momento o resultado deve estar impresso na fita, lendo de trás para frente, 1100, equivalente a 12 decimal (Figura 3).

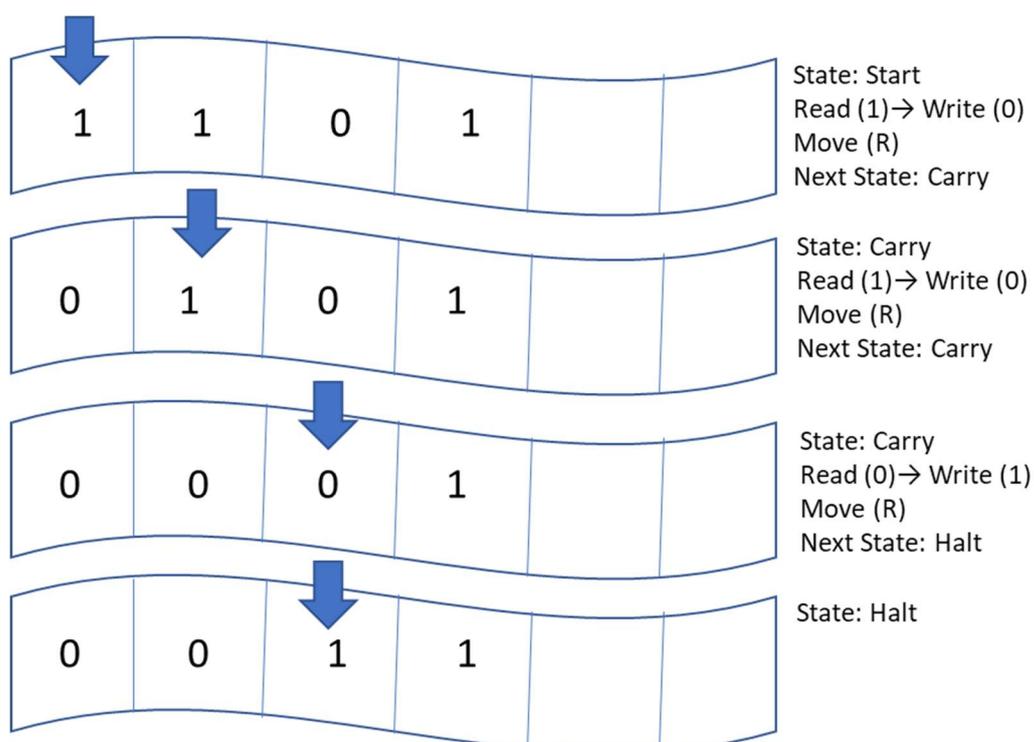


Figura 3: Exemplo de operações de realizadas por uma TM ao incrementar o número binário 1011 em uma unidade. Fonte: elaboração do autor.

⁹ Código para este exemplo disponível no Apêndice A.

Com isto Turing pretendia estabelecer uma definição precisa do conceito de computação de forma que todas as operações realizadas na matemática poderiam ser descritas como uma série de operações simples, tais quais a da TM.

Seu próximo passo, portanto, foi imaginar uma máquina que pudesse emular outras, isto é, uma máquina que tomasse além dos dados, também a tabela de operações através da fita. A primeira parte da fita conteria uma descrição da TM a ser emulada (o conjunto de estados e regras de transição), e a segunda parte da fita contém as entradas a serem processadas. A isto chamou de máquina computacional universal, que ficou conhecido como Máquina de Turing Universal (UTM).

Turing buscava uma forma de abordar um problema que estava gerando bastante confusão na matemática do início do século XX: o problema da autorreferência. Um exemplo disto é o paradoxo do mentiroso (BEALL, GLANZBERG, *et al.*, 2020), que já era conhecido desde a Grécia antiga, e pode ser sumarizado pela sentença:

Esta proposição é falsa.

Neste caso, se a afirmação é falsa, seu valor lógico é verdadeiro. Porém se for verdadeira, tem valor lógico falso. Este tipo de problema aparece com força na matemática a partir da comunicação entre o famoso matemático e lógico alemão Gottlob Frege (1848 – 1925) e o então jovem filósofo britânico Bertrand Russel (1872 – 1970), no qual este apresentou o que ficou conhecido como *Paradoxo de Russel* (DAVIS, 2001), que pode ser formulado da seguinte maneira:

Seja Z o conjunto de todos os conjuntos que não contém a si mesmos.

$$Z = \{x \mid x \notin x\}$$

Pergunta: A está contido em A?

Se este conjunto contém a si mesmo, então, por definição, não deveria. E se não, então, por definição, deveria. Em outras palavras, usando a ilustração do barbeiro (IRVINE, DEUTSCH, 2021):

Suponha que exista uma cidade com um barbeiro que faz a barba de todos e somente daqueles homens que não fazem a própria barba. Quem faz a barba do barbeiro?

Se ele fizer a própria barba, então ele não deveria fazer (pois ele só faz a barba dos homens que não fazem a própria barba), e se ele não fizer a própria barba, então

ele deveria fazer (pois ele faz a barba de todos os homens que não fazem a própria barba). Russell utilizou esse paradoxo para ilustrar a inconsistência da teoria ingênua dos conjuntos, que permite que qualquer coleção definível seja considerada um conjunto. De acordo com a teoria ingênua dos conjuntos, qualquer condição (propriedade definível) deveria determinar um conjunto de coisas que satisfazem essa condição. O Paradoxo do Barbeiro demonstra que essa suposição leva a uma contradição.

Alan Turing procurava demonstrar que todo sistema matemático baseado em regras está sujeito a contradições, e ele usou o conceito de UTM para alcançar esse objetivo. A UTM consiste em uma tabela de descrição, que chamaremos de *tabela operacional*, que recebe como entrada outra tabela de descrição, denominada aqui de *tabela objeto*, e os dados a serem operados. A função da UTM é fornecer como saída uma resposta do tipo verdadeiro ou falso, isto é, um *problema de decisão*.

Quadro 3: Implementação em pseudocódigo do que seria um programa que avalia se o programa P com a entrada I termina ou não. Ao final o programa toma a si mesmo como objeto. Fonte: Elaboração do autor.

```

Define Halt(P, I)
    // This is a hypothetical procedure that determines if program P
    // halts on input I. We don't know how to implement this,
    // and Turing says we can't
    // For the purposes of this illustration, assume it correctly
    // returns 'true' if the program halts and 'false' otherwise.
End Define

Define Test(P, I)
    If Halt(P, I) Then
        While True Do
            // Do nothing, loop forever
        End While
    Else
        Return
        // Halt
    End If
End Define
Test(Test,Test)

```

A resposta será "verdadeiro" se a tabela objeto implementar uma série de operações que eventualmente chegam ao fim (halt) quando executadas. Por outro lado, se a tabela objeto executar indefinidamente, como em um loop repetindo as mesmas operações ciclicamente, a resposta será "falso". A UTM toma, então, essa resposta e executa o oposto do que a tabela objeto faria: se a resposta for "verdadeiro", a tabela operacional entra em um loop, e se a resposta for que o programa objeto entra em um loop, a tabela operacional para. Usando o argumento

por absurdo, Turing demonstra que tal tabela operacional não pode existir pois, ao submeter ela como sua própria tabela objeto, se tal tabela chega ao estado de *halt*, ela entra em loop, mas se ela entra em loop, então deve parar. Uma contradição (Quadro 3).

O Problema de Parada, ou *Halting Problem*, revela uma característica importante sobre a natureza da matemática e da computação: nem todas as perguntas ou problemas que podem ser claramente definidos podem ser definitivamente respondidos ou resolvidos por meio de um algoritmo mecânico, independentemente de quão sofisticado esse algoritmo seja. Antes da formalização do Problema de Parada, acreditava-se que para cada problema matemático bem-definido existia um algoritmo que poderia decidir a resposta. Existem, porém, problemas bem-definidos, como "esse programa vai parar com essa entrada?", para os quais um algoritmo que possa decidir a resposta em todos os casos não pode ser criado.

O Problema de Parada tem relações com os Teoremas da Incompletude de Gödel, os quais afirmam que em qualquer sistema matemático formal consistente capaz de expressar certas expressões aritméticas, haverá proposições verdadeiras que não podem ser provadas dentro desse sistema¹⁰. No caso do Problema de Parada, a pergunta indecidível é se um determinado programa para ou não.

Outra implicação importante do Problema de Parada é que ele demonstra que até mesmo sistemas determinísticos, como máquinas de Turing ou programas de computador, podem exibir um comportamento intrinsecamente imprevisível. Essa imprevisibilidade não é causada por aleatoriedade ou caos, mas sim pela estrutura lógica da computação em si. Isso tem implicações profundas para a compreensão de sistemas complexos em matemática e ciência. Além disso, o Problema de Parada tem relevância na teoria da complexidade computacional. Ele mostra que existem problemas para os quais não podemos antecipar os recursos computacionais necessários para resolvê-los. Isso levou ao desenvolvimento da teoria da NP-completude, que classifica uma grande classe de problemas como sendo computacionalmente intratáveis.

¹⁰ Esta maneira de descrever é bastante genérica e, portanto, pouco acurada, porém não é a proposta aqui investigar o escopo de validade do primeiro teorema de incompletude, apenas apontar que os resultados obtidos por Turing são equivalente aos de Gödel (HOFSTADTER, 1999).

4 TEORIA ALGORÍTMICA DA INFORMAÇÃO

Pode-se dizer, portanto, que aquele que age de forma perfeita é como um excelente Geômetra que sabe como encontrar a melhor solução para um problema; como um bom arquiteto que utiliza sua localização e os recursos destinados à construção da maneira mais vantajosa, não deixando nada que seja distoante ou que não exiba a beleza de que é capaz; como um bom proprietário que utiliza sua propriedade de forma a não deixar nada inculto ou estéril; como um engenheiro habilidoso que produz de maneira menos difícil possível; e como um autor inteligente que condensa a maior quantidade de realidade no menor espaço possível.

Leibniz

A Teoria da Informação Algorítmica (*Algorithmic Information Theory* – AIT), um dos referenciais teóricos para o desenvolvimento de aprendizado de máquina (*Machine Learning*) e predição, deve seu desenvolvimento às contribuições de pioneiros como Ray Solomonoff (1926 – 2009), Andrey Kolmogorov (1903 – 1987) e Gregory Chaitin¹¹ (nascido em 1947).

Nos anos sessentas do século passado, Solomonoff introduziu a o que ficou conhecido como *Indução de Solomonoff* (SOLOMONOFF, 1964a, b), um modelo teórico que ganhou reconhecimento como uma abordagem ideal para a predição, embora sua aplicação computacional seja inviável. Kolmogorov e Chaitin, com seus trabalhos independentes, mas relacionados, adicionaram elementos significativos: Kolmogorov introduziu uma medida de complexidade, a qual ficou conhecida como Complexidade de Kolmogorov (*Kolmogorov Complexity* – KC), que quantifica os recursos computacionais necessários para especificar um objeto, como uma sequência de dados. Chaitin estendeu ainda mais essas ideias, contribuindo para a compreensão da aleatoriedade inerente e da incomputabilidade, associada à KC, com o conceito do número ômega.

¹¹ Professor visitante estrangeiro da Universidade Federal do Rio de Janeiro – UFRJ por três vezes, para o programa de pós-graduação em História das Ciências e das Técnicas e Epistemologia – HCTE.

Apesar da inerente intractabilidade computacional, essas teorias moldaram extensivamente a maneira como compreendemos e concebemos algoritmos de aproximação práticos para aprendizado de máquina e predição. Este capítulo apresenta uma revisão teórica abrangente da AIT com foco na Indução de Solomonoff, discutindo seus fundamentos teóricos, implicações e limitações. Também exploramos as contribuições de Kolmogorov e Chaitin para esse domínio, refletindo sobre como seu trabalho enriqueceu o panorama da Ciência da Computação teórica.

4.1 Problema da indução em Solomonoff

A Indução de Solomonoff oferece uma formalização matematicamente rigorosa e computacional da epistemologia bayesiana, apresentando um modelo teórico de referência para a predição (SOLOMONOFF, 1964a, b). Embora a inviabilidade computacional do método de Solomonoff impossibilite sua aplicação direta na prática, ele serve como um ponto de referência fundamental para a predição, aprendizado de máquina e inteligência artificial, estabelecendo uma métrica que pode ser aplicável ao desenvolvimento de algoritmos de aproximação mais eficientes na prática. A ideia básica da Indução de Solomonoff é que, ao prever o futuro com base no passado, deve-se considerar todas as possíveis hipóteses que podem explicar o que já se sabe.

O método de Solomonoff segue o paradigma bayesiano, enxergando a tarefa de predição como um exercício de inferência bayesiana (SHEN, 1999, VITANYI, 2010). Partindo de dados passados, o método formula todas as possíveis explicações e atribui probabilidades a priori a essas hipóteses. O Teorema de Bayes é um teorema fundamental na teoria da probabilidade e estatística que descreve como atualizar as probabilidades das hipóteses ao receber evidências. Ela recebe esse nome em homenagem a Thomas Bayes (1701-1761), que foi o responsável pela primeira formulação matemática desse problema.

Assim, em um problema de inferência indutiva, o objetivo é determinar qual ou quais hipóteses, de um conjunto de hipóteses $H=h_1, h_2, \dots$, são as mais prováveis de serem responsáveis por um conjunto de dados observados $D=x_1, x_2, \dots$. Nesse contexto, o Teorema de Bayes, oferece uma formulação para esse problema.

Através da aplicação da equação (Eq. 6), novas evidências atualizam a probabilidade de um evento a partir do conhecimento *a priori* (ou a crença inicial na ocorrência de um evento). Assim, uma limitação para o uso do Teorema de Bayes vem justamente de não se conhecer as probabilidades a priori, isto é, antes de se obter qualquer dado a respeito do problema em mãos.

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)}$$

Eq. 6: Teorema ou Regra de Bayes. Adaptado de (THRUN, BAXTER, *et al.*, 1997)

Onde,

- $P(h|D)$ é a probabilidade posterior, representando a probabilidade da hipótese h ser verdadeira dado o conjunto de dados D .
- $P(D|h)$ representa a probabilidade de observar o conjunto de dados D assumindo que a hipótese h seja verdadeira.
- $P(h)$ é a probabilidade a priori, representando a crença inicial na hipótese h antes de observar os dados.
- $P(D)$ representa a probabilidade total de observar o conjunto de dados D .

Esse processo é comumente percebido como subjetivo, pois reflete as crenças ou suposições iniciais de cada indivíduo. No entanto, a busca por uma abordagem mais objetiva e fundamentada teoricamente na determinação dessas probabilidades a priori levou ao desenvolvimento da AIT. Para tanto, Solomonoff estabeleceu a relação com o trabalho de Turing, propondo que um programa com menor extensão, ou seja, um conjunto mais compacto de operações, teria uma maior probabilidade de explicar um determinado fenômeno. Em outras palavras, um programa de computador que produz os dados observados como saída ao ser executado em uma máquina de Turing universal.

A probabilidade a priori de uma hipótese é inversamente proporcional à sua complexidade, sendo a simplicidade associada a uma maior probabilidade. Esse viés em direção à simplicidade surge dos princípios da AIT, em que a complexidade de uma hipótese corresponde ao comprimento do menor programa de máquina de Turing capaz de reproduzir os dados observados. (HUTTER, 2009).

Ao dar mais peso a hipóteses mais simples, o método de Solomonoff incorpora o espírito da Navalha de Ockham. A *simplicidade* de uma hipótese corresponde ao

comprimento do programa mais curto: quanto mais curto o programa, mais simples a hipótese.

A Navalha de Ockham é um princípio filosófico nomeado em homenagem a Guilherme de Ockham (1285 – 1349 EC), lógico e frade franciscano. Ele postula que, diante de um conjunto de hipóteses concorrentes que explicam suficientemente os fenômenos observados, a mais simples deve ser selecionada. O termo "mais simples" geralmente se refere à hipótese com menos pressuposições. No contexto de uma distribuição a priori sobre hipóteses, isso implica em atribuir uma probabilidade a priori mais alta às hipóteses mais simples e uma probabilidade decrescente à medida que as hipóteses se tornam mais complexas. Essa abordagem reflete a preferência por soluções mais simples e parcimoniosas, ou, talvez, mais elegantes (MARKS, CHAITIN, 2021).

Uma vez que um método tenha sido estabelecido para se classificar as hipóteses em termos de probabilidade de explicar um fenômeno, é necessário criar uma lista contendo todas as possíveis hipóteses capazes de explicar os dados observados. Esse procedimento está em consonância com um princípio filosófico atribuído a Epicuro (342? AEC - 270 AEC), que preconiza a retenção de todas as hipóteses consistentes com os dados. A fim de evitar o descarte completo de hipóteses consistentes, é fundamental atribuir a cada uma delas uma probabilidade a priori que não seja nula. Na notação da Regra de Bayes, isso é expresso como $\forall h : P(h) > 0$, todas as hipóteses h cuja probabilidade a priori de h seja maior que zero serão mantidas.

Esse princípio está alinhado com a prática científica de manter a mente aberta e revisar crenças à luz de novas evidências. Ele previne o risco de descartar hipóteses verdadeiras que, à primeira vista, podem parecer implausíveis ou que ainda não possuem evidências substanciais para sustentá-las. A implicação desse princípio filosófico na prática é que mesmo hipóteses que pareçam altamente complexas ou improváveis com base nos dados atuais ainda são mantidas em consideração, embora com baixas probabilidades. À medida que mais dados são observados, essas probabilidades são atualizadas. Assim, se dados futuros começarem a favorecer uma hipótese anteriormente improvável, sua probabilidade poderá aumentar de acordo.

Como ilustração, considere um cenário em que um indivíduo interage com uma plataforma de mídia social. À medida que o usuário se engaja com a plataforma ao

longo do tempo, ele começa a observar um padrão particular em seu *feed*: o algoritmo parece servir conteúdo em uma sequência consistente – primeiro, uma postagem de um amigo próximo, depois uma postagem de um grupo que ele segue e, finalmente, um anúncio. Esse padrão discernido informa a expectativa do usuário sobre que tipo de postagem ele encontrará em seguida ao rolar o feed. Isso é indicativo de uma forma de raciocínio indutivo, em que um padrão observável é generalizado para prever instâncias futuras. Nesse cenário, o usuário espera encontrar postagens na mesma sequência – amigo, grupo, anúncio - à medida que continua a interagir com a plataforma.

No entanto, suponha que haja uma alteração repentina e não anunciada no algoritmo da plataforma. Consequentemente, a sequência de postagens muda para – amigo, amigo, anúncio, grupo. Essa mudança abrupta interrompe o padrão previamente discernido, resultando na expectativa do usuário estar incorreta.

Nesse contexto, a Indução de Solomonoff apresenta um modelo mais sofisticado para prever a próxima postagem no feed. Ela assume um a priori universal e calcula a probabilidade posterior sobre todas as hipóteses ou padrões possíveis que podem descrever a sequência de postagens. Embora atribua probabilidades mais altas a padrões mais simples - em consonância com o princípio da navalha de Ockham - ela ainda considera todos os padrões dentro do espaço de hipóteses. No caso do feed de notícias alterado do usuário, a Indução de Solomonoff provavelmente iria prever que a próxima postagem seria de um grupo, conforme o padrão observado com mais frequência. No entanto, também consideraria a possibilidade de a postagem ser um anúncio ou até mesmo de outro amigo, atribuindo essas possibilidades com probabilidades menores.

A limitação fundamental para a direta aplicação desta abordagem se dá por ela requer considerar todas as possíveis hipóteses, o que é computacionalmente inviável. Outra questão é que ela requer atribuir probabilidades às hipóteses com base em sua complexidade, mas a complexidade de uma hipótese pode ser incomputável devido ao Problema de Parada - nem sempre é possível determinar quanto tempo um programa de computador levará para parar e, portanto, não é possível determinar sempre a complexidade de uma hipótese.

Apesar dessas limitações, a Indução de Solomonoff apresenta semelhanças com o campo da criptografia, especialmente no contexto da compressão de dados.

Quando os dados são reduzidos à sua representação mais simples, correspondente ao programa mais curto capaz de gerar esses dados, muitas vezes eles se tornam ininteligíveis sem se executar o programa para reverter o processo de compressão. Isso se assemelha à ofuscação de informações na criptografia, em que os dados permanecem ininteligíveis sem a chave de descriptação apropriada. No entanto, no contexto da AIT – a teoria fundamental por trás da Indução de Solomonoff – essas analogias criptográficas surgem não de uma intenção explícita de proteger os dados, mas sim do processo intrínseco de compressão de dados.

Para elucidar os conceitos de complexidade e aleatoriedade dentro desse contexto, retomemos o exemplo da moeda do Capítulo 3. Considere a geração de uma sequência binária por meio de uma série de lançamentos de uma moeda justa. Façamos que um lançamento seja descrito como “0” quando o resultado for “cara” e “1” quando o resultado for “coroa”. Após um certo número n de lançamentos obteremos uma cadeia *string* de 1s e 0s com o comprimento n . Logo, a probabilidade de uma string específica é dado por 2^{-n} . Entretanto, uma sequência composta apenas por 1s possui um padrão discernível, o que a distingue da maioria das outras sequências que pareceriam aleatórias e indistinguíveis. Intuitivamente, parece mais provável que um processo determinístico de certa simplicidade, em vez de um processo aleatório, tenha produzido tal padrão.

Em conclusão, a Indução de Solomonoff oferece um quadro probabilístico abrangente que equilibra simplicidade e complexidade na previsão de sequências. Embora limitações computacionais e teóricas dificultem sua aplicação direta, seus princípios continuam a informar e moldar nossa compreensão do reconhecimento de padrões, compressão de dados e complexidade no campo da teoria da informação. Ao elucidar as conexões entre aleatoriedade e complexidade, a Indução de Solomonoff oferece uma compreensão profunda sobre a natureza dos dados, das previsões e da busca pela simplicidade no vasto panorama de hipóteses potenciais.

4.2 Complexidade de Kolmogorov

A Complexidade de Kolmogorov (KC) é uma ferramenta teórica abrangente com diversas aplicações em diferentes campos, como ciência da computação, física, estatística e ciência cognitiva. Ela desempenha um papel fundamental em discussões sobre informação, aleatoriedade e complexidade. Essa medida oferece uma maneira de quantificar a complexidade intrínseca de uma sequência de dados, medindo o tamanho do programa mais curto que pode produzir essa sequência em uma UTM. Com base nessa definição, a KC nos permite explorar a natureza fundamental da informação, examinando a existência de padrões, estruturas ou regularidades em uma sequência de dados. Essa ferramenta tem sido amplamente utilizada para investigar questões em áreas como a compressão de dados, a teoria da informação, a criptografia, a aprendizagem de máquina e a teoria da complexidade computacional.



Figura 4: Imagem de uma torre de legos com padrão mais complexo de cores gerado através da tecnologia *Stable Diffusion* pelo *prompt*: a lego tower of a few blocks of alternating colors.

Vamos tomar como exemplo uma construção simples de Lego (Figura 5), como uma torre feita de quatro blocos de cores uniformes empilhados. O conjunto de instruções necessário para replicar essa construção seria conciso e direto, indicando algo como: empilhe verticalmente quatro blocos vermelhos. Nesse cenário, a KC da torre de Lego é baixa, refletindo sua simplicidade. À medida que a complexidade da

construção de Lego aumenta, como uma torre com um padrão de cores alternadas, a KC também aumenta. As instruções de replicação se tornam mais detalhadas, possivelmente indicando: Comece com um bloco vermelho, depois empilhe um bloco azul em cima, seguido por outro bloco vermelho e, por fim, termine com um bloco azul.

No caso de uma construção completamente aleatória, onde não há um padrão ou regra discernível que governe a disposição dos blocos de Lego, a KC atinge seu máximo. As instruções precisariam especificar a posição, orientação e cor de cada bloco individualmente. Por exemplo: Coloque um bloco azul deitado, depois coloque um bloco vermelho em pé à esquerda, um bloco verde deitado no centro e um bloco amarelo em pé à direita. Assim, por meio dessa analogia, fica evidente que a KC de um objeto é essencialmente uma medida de sua complexidade estrutural. Quanto mais complexo ou aleatório o objeto, mais longas serão as instruções necessárias para sua replicação (CHAITIN, 1974) e, portanto, maior será sua KC.

Essa medida, portanto, representa a quantidade de informações necessárias para descrever uma sequência de forma completa e precisa. Formalmente, a KC de uma sequência x , denotada por $K(x)$, é definida como o comprimento do programa binário mais curto p possível que, quando executado em uma UTM U , produz a sequência x como saída e, em seguida, termina:

$$K(x) = \min_{p: U(p)=x} |p|,$$

onde $|p|$ é o comprimento de p e a minimização é feita sobre todos os programas binários p que geram a sequência x quando executados em U .

Surge dessa maneira uma conexão com a compressão de dados pois, se pudéssemos calcular a KC para qualquer conjunto de dados, seríamos capazes de identificar o programa mais curto para reproduzir esses dados. Isso permitiria a compressão ótima dos dados, encontrando a sequência binária mais curta que poderia ser utilizada para recriar os dados originais. Uma sequência é considerada incompressível se não tiver uma descrição mais curta do que ela mesma. Se uma sequência for aleatória (sem nenhum padrão discernível), sua KC é aproximadamente equivalente ao seu comprimento. Para uma sequência x , isso implica:

$$K(x) = |x|, \text{ se } x \text{ for incompressível.}$$

No entanto, é importante ressaltar que não podemos determinar a KC de uma sequência arbitrária por meio de um algoritmo universal. Essa impossibilidade deriva diretamente Problema de Parada, que afirma que não existe uma solução algorítmica capaz de prever, de forma geral, se um programa irá terminar sua execução ou não. Portanto, encontrar o programa mais curto que reproduz qualquer conjunto de dados de maneira eficiente é uma tarefa não computável. Formalmente, não existe um programa H tal que para todas as sequências x , $H(x) = K(x)$.

4.3 Descrição de Comprimento Mínimo

A origem do princípio da Descrição de Comprimento Mínimo (MDL) tem origem no trabalho seminal de Jorma Rissanen (1932 - 2020), uma figura influente no campo da teoria da informação e estatística. Em seu artigo de 1978, *Modeling by shortest data description*, Rissanen introduziu o princípio de que o modelo mais adequado para um determinado conjunto de dados é aquele que fornece a descrição mais curta dos dados (RISSANEN, 1978). Essa perspectiva ofereceu um critério quantificável que facilita a seleção do modelo mais adequado para qualquer conjunto de dados, abrindo espaço para avanços subsequentes em compressão de dados, reconhecimento de padrões e modelagem estatística.

O MDL *idealizado*, por sua vez, é um descendente direto da teoria da inferência indutiva de Ray Solomonoff, que propôs que o modelo mais eficaz para um conjunto de dados específico é aquele representado pelo programa mais curto que pode reproduzir o conjunto de dados quando executado em uma UTM (SOLOMONOFF, 1964b, a). O uso do termo “idealizado”, reflete que, em geral, chegar a este programa é um procedimento incomputável. No entanto, é crucial notar que, enquanto a teoria de Solomonoff está focada em encontrar o programa mais curto possível que reproduza os dados, o objetivo principal do princípio MDL é um pouco diferente. Ele busca o modelo que minimiza o comprimento total da descrição, que inclui tanto a complexidade do modelo em si quanto o comprimento da descrição dos dados quando codificados usando esse modelo. Portanto, embora o princípio MDL idealizado e a teoria de Solomonoff se originem do campo da Teoria da Informação

Algorítmica e compartilhem semelhanças, eles abordam aspectos ligeiramente diferentes da descrição de dados (POLAND, HUTTER, 2005).

Incorporado no princípio MDL está um importante compromisso entre a complexidade do modelo e seu poder explicativo sobre os dados. Isso pode ser exemplificado com um conjunto de dados simples que compreende as notas dos alunos e as horas que eles estudam por semana. Embora um polinômio de grau superior possa se ajustar a cada ponto de dados com precisão, ele aumenta a complexidade do modelo. Um modelo de regressão linear mais simples pode não levar em conta todos os pontos de dados com precisão, mas provavelmente terá um comprimento de descrição geral menor uma vez que os erros são incluídos. Portanto, de acordo com o MDL, o modelo mais simples pode ser uma escolha preferível evidenciando o equilíbrio necessário entre complexidade e precisão.

O princípio MDL e sua busca por simplicidade também têm ressonâncias com o conceito de KC. Sob esta ótica, conjuntos de dados simples têm um comprimento de descrição mais curto e, portanto, a KC é menor, enquanto conjuntos de dados complexos ou aleatórios, que resistem à compressão, têm uma KC maior ou igual ao comprimento dos próprios dados. O objetivo de minimizar o comprimento da descrição, portanto, se alinha com a redução da KC (POLAND, HUTTER, 2005).

No cerne do princípio MDL está a formalização da Navalha de Ockham, afirmando que a explicação ou modelo mais simples, que faz as menores suposições, é frequentemente o melhor (STERKENBURG, 2016). Ele estimula o uso de modelos simples que podem descrever adequadamente os dados e a seleção de classificadores *mais simples* em detrimento dos mais complexos. Isso visa identificar a representação mais reduzida e essencial de todos os membros de uma categoria, com as variações entre os padrões individuais consideradas como *ruído*. Simplificando adequadamente, esse ruído pode ser reduzido, e o *signal* fundamental pode ser retido.

A aplicação prática do princípio MDL oferece duas metodologias distintas de previsão e tomada de decisão: um método *seguro* e um método *arriscado*. Um modelo sob o método seguro dá uma impressão precisa do erro de previsão para dados futuros, mesmo que o modelo simplifique o processo subjacente. Em contraste, o método arriscado não oferece essa garantia, mas pode fornecer resultados úteis.

Para ilustrar essas metodologias, considere o domínio da previsão do tempo. Modelos categorizados como seguros objetivam a previsão de padrões meteorológicos médios, mantendo um alto grau de precisão a longo prazo, mesmo quando o processo subjacente é simplificado. Por outro lado, os modelos arriscados tentam prever condições climáticas diárias específicas, oferecendo detalhes valiosos para situações particulares. Contudo, esses últimos não proporcionam a garantia de prever com precisão os eventos futuros. Você poderá acabar carregando o dia todo um guarda-chuvas sem que de fato a chuva ocorra.

Em certas áreas, o método seguro é preferível, por exemplo, um médico ao diagnosticar um paciente opta por um diagnóstico que se encaixa estatisticamente nos sintomas, mesmo que isso simplifique a complexidade da condição real do paciente. Esta abordagem visa minimizar o risco de danos graves por meio de um diagnóstico errado, mas pode falhar em descobrir uma doença mais rara. Por outro lado, um *trader* de ações pode favorecer o método arriscado: pode escolher um modelo que prevê movimentos diários de ações, embora com uma chance maior de erros, em vez de um que prevê com segurança as tendências de longo prazo. Aqui, o potencial de um lucro substancial pode superar o risco das imprecisões diárias.

A interação entre o princípio MDL e a probabilidade, quando examinada através da lente da codificação da probabilidade, apresenta uma perspectiva original. Os métodos estatísticos tradicionais geralmente dependem das interpretações da probabilidade como a frequência de ocorrência de um evento (interpretação frequentista) ou a força de crença sobre um evento (interpretação bayesiana). Por exemplo, em uma análise frequentista, um modelo pode ser escolhido com base na frequência com que um determinado conjunto de dados ou padrão é observado em uma grande amostra. Em uma análise bayesiana, a escolha do modelo pode ser orientada pelo grau de crença ou evidência que apoia uma determinada hipótese (HUTTER, 2003).

Por outro lado, o princípio MDL se concentra no conteúdo de informação de um conjunto de dados. Em vez de se preocupar com a frequência de ocorrência ou crenças sobre um evento, o princípio MDL busca o modelo que fornece a descrição mais sucinta do conjunto de dados. Este modelo é visto como aquele que melhor representa os dados.

Por exemplo, considere um conjunto de dados que representa a precipitação anual em uma região. Uma abordagem frequentista tradicional pode escolher um modelo com base na frequência com que uma determinada quantidade de chuva é observada, enquanto uma abordagem bayesiana poderia selecionar um modelo baseado em estudos meteorológicos prévios sobre os padrões de chuva. Em contraste, o princípio MDL visaria encontrar um modelo que oferece a descrição mais sucinta ou compacta dos dados de chuva, independentemente da frequência de quantidades específicas de chuva ou crenças prévias.

Esta ênfase no comprimento da descrição desloca o foco da probabilidade como uma propriedade inerente do sistema ou como uma crença subjetiva, e em vez disso, enquadra a probabilidade em termos de codificação de informação. Isso significa que eventos que são comuns – e, portanto, mais previsíveis – podem ser representados por códigos mais curtos, enquanto eventos mais raros exigirão códigos mais longos.

Considere o exemplo de uma equação física, que pode ser vista como um modelo que descreve de forma concisa um conjunto de dados, ou seja, o comportamento observado de um sistema. A bem conhecida lei da gravitação universal de Newton,

$$F = G \frac{m_1 m_2}{r^2}$$

serve como um exemplo de tal modelo. Apesar de sua simplicidade, esta equação prevê com precisão uma ampla variedade de fenômenos, desde a anedótica maçã caindo no chão até a órbita da lua, servindo assim como um código eficiente.

No entanto, quando se trata de fenômenos que são mais raros ou complexos, como o comportamento de partículas na mecânica quântica, as equações correspondentes se tornam substancialmente mais complexos. Este exemplo ilustra como o princípio do MDL se desvia dos métodos estatísticos convencionais, enfatizando o conteúdo de informação ou comprimento da descrição sobre a frequência de ocorrência ou crença.

Ainda assim, mesmo a equação física mais compacta não captura tudo que é aleatório sobre o sistema. A aleatoriedade ou incerteza inerente ao sistema se manifesta como erro ou ruído na representação dos dados pelo modelo. De acordo com o princípio do MDL, esse erro contribui para o comprimento total da descrição do modelo e dos dados.

Em sua busca para construir modelos que minimizem o comprimento total da descrição, o princípio do MDL leva em conta tanto a complexidade do modelo quanto a capacidade do modelo de codificar os dados fornecidos. Esta abordagem preenche a lacuna entre a compressão de dados e a probabilidade, oferecendo um novo ponto de vista sobre a seleção de modelos e a análise de dados. Consequentemente, a melhor explicação dos dados é identificada como aquela que fornece a descrição mais concisa, não muito diferente das equações físicas mais *elegantes* e concisas.

4.4 Número ômega

Solomonoff descobriu um teorema (SOLOMONOFF, 1960) que afirma que entre todos os algoritmos capazes de decodificar strings a partir de suas descrições ou códigos, existe um que é ótimo. Este teorema nos assegura que, embora o valor absoluto da KC dependa da escolha da UTM, essa dependência é limitada por uma constante. Consequentemente, para cadeias de caracteres suficientemente longas, a escolha da máquina universal não é significativa – é por isso que podemos falar sobre a KC de uma cadeia de caracteres de forma absoluta, mesmo que tecnicamente dependa da escolha da máquina. Esse teorema foi descoberto independentemente por Kolmogorov e Chaitin (CHAITIN, 1975).

O número ômega (ou *número* Ω), conceito que foi desenvolvido por Gregory Chaitin, é um número real que encapsula a probabilidade de um programa de computador gerado aleatoriamente terminar. Embora o conceito do número Ω pareça singular, na realidade existem infinitas variações deste e cada variante corresponde aos diferentes métodos pelos quais os programas de computador podem ser codificados. Como resultado, deve-se entender que a referência geralmente é feita a uma família de números Ω , e não apenas a uma única entidade. Além disso, a dependência de Ω do método de codificação do programa motiva o uso do termo *construção de Chaitin* ao discutir Ω de maneira independente da codificação.

Cada instância do número Ω é um número real, entre 0 e 1, *normal* e *transcendente*, representando duas características-chave na teoria dos números. A normalidade implica que, na expansão infinita da representação decimal de Ω , cada dígito e combinação de dígitos aparecem com a mesma frequência, possuem a

mesma densidade. A transcendentalidade indica que o número não pode ser uma raiz de nenhuma equação polinomial não nula com coeficientes racionais.

Cada UTM possui o seu número Ω : Se considerarmos o conjunto de todos os programas binários que param de executar quando rodados em uma máquina de Turing específica e os listarmos em ordem lexicográfica, o n -ésimo programa contribui com 2^{-n} para Ω (KLEENE, 1988). Cabe destacar que *cada* número Ω , embora bem definido, é não computável, o que implica que não existe um algoritmo capaz de calcular seus dígitos exatamente. Essa não computabilidade é uma consequência direta do Problema de Parada. Isto é, tal como a KC, se fosse possível calcular o número Ω seria possível por consequência imediata resolver o Problema de Parada, mas este, sabemos, não ter solução: O n -ésimo dígito de Ω indica se o n -ésimo programa, na ordem lexicográfica, para ou não de executar, mas não há um padrão no comportamento de parada das máquinas de Turing.

Em vista disso, decorre que cada número Ω exibe uma propriedade chamada aleatoriedade de Martin-Löf (DOTY, 2006). Per Erik Rutger Martin-Löf é um lógico e matemático sueco, nascido em 1942, que foi orientando de Kolmogorov.

A aleatoriedade de Martin-Löf implica que nenhum algoritmo pode prever ou adivinhar confiavelmente os dígitos do número, ou, ao menos, não possui um desempenho melhor que um processo meramente aleatório. É importante enfatizar que isso não se deve a uma falta de precisão nos algoritmos, mas sim a uma propriedade intrínseca do próprio número, destacando, portanto, a complexidade inerente desses números e sua resistência à análise computacional.

Ao compreendermos os limites fundamentais impostos pelos desafios computacionais e a natureza complexa e imprevisível da aleatoriedade, somos instigados a reavaliar nossas abordagens e a explorar novas perspectivas na busca por um entendimento mais abrangente da realidade. Nesse contexto, a AIT surge como uma ferramenta teórica que nos permite investigar e quantificar a complexidade dos sistemas de informação. Ao debruçarmo-nos sobre os enigmas da incomputabilidade, somos conduzidos a uma exploração acerca dos limites do conhecimento humano, desvelando uma visão mais refinada sobre a complexidade e aleatoriedade presentes no processamento de informações.

5 REFLEXÕES FINAIS

Problemas incomputáveis, como o Problema de Parada, surgem quando não é possível encontrar um algoritmo universal que resolva todas as instâncias desses problemas. Essa característica ressalta a complexidade intrínseca de certas tarefas computacionais e impõe limites à resolução sistemática de problemas. Embora tenhamos a tendência de procurar por padrões em todas as áreas, não está claro se os padrões que observamos, que fundamentam nosso senso comum e o conhecimento científico, são uma ocorrência comum ou uma exceção. Ao dizer que uma sequência é considerada aleatória ou incompressível quando possui uma alta KC, estamos afirmando que ela não pode ser descrita de forma mais concisa ou comprimida. Essa noção de aleatoriedade sugere que dados verdadeiramente aleatórios não apresentam nenhum padrão ou estrutura discernível que possa ser explorado para reduzir o tamanho de sua descrição. Mais que uma limitação epistemológica, este fato pode apontar para uma problemática ontológica, o que dizer de um fenômeno cuja melhor descrição é, no final das contas, o próprio fenômeno?

De fato, a constatação de que existem problemas incomputáveis e que a aleatoriedade genuína não apresenta padrões discerníveis levanta questões filosóficas e ontológicas fascinantes. Essas questões nos levam a refletir sobre a natureza da realidade e a compreensão que temos do universo. A existência de problemas incomputáveis implica que há limites fundamentais para nossa capacidade de compreender e prever certos aspectos do mundo. Isso exorciza o demônio de Laplace desafiando a visão determinística clássica, que sugeria que, em princípio, tudo pode ser conhecido e calculado. A complexidade e a aleatoriedade demonstram que há fenômenos e padrões que estão além do alcance da computação, o que nos leva a questionar se a realidade não seria, em si, intrinsecamente indeterminada ou imprevisível em certos aspectos.

A aleatoriedade também nos faz refletir sobre a própria natureza do conhecimento e da informação. Se dados verdadeiramente aleatórios não apresentam padrões discerníveis, isso sugere que há um limite para nossa capacidade cognitiva em compreender plenamente o universo. Isso nos convida a repensar a relação entre a descrição da realidade e a própria realidade,

questionando se a melhor descrição da realidade é apenas uma aproximação limitada ou se a realidade em si é intrinsecamente complexa e indescritível em sua totalidade.

O conceito de complexidade e AIT nos leva, também, a considerar como alguns sistemas de informação na natureza podem representar dados *comprimidos* e em como isso afeta nosso acesso a compreensão de fenômenos naturais. O DNA é um exemplo paradigmático: as informações codificadas nos pares de bases de nucleotídeos no DNA não ditam explicitamente como criar outro organismo. Em vez disso, estabelecem uma planta baixa, ou um conjunto de instruções, para processos biológicos. Nesse contexto, as relações entre os dados não são imediatamente aparentes, pois um gene não apenas codifica proteínas, acima disso, sua localização (ou *locus*) no DNA pode ter um efeito regulador na expressão de outros genes. Isso ilustra que o DNA não é um sistema simples de armazenamento unidimensional de dados. Em vez disso, é um sistema complexo e multidimensional, onde os dados (genes) estão interconectados, influenciando a expressão uns dos outros – uma manifestação da complexidade e teoria da informação na biologia.

Richard Dawkins inicia um capítulo de seu livro, *The Blind Watchmaker*, com a frase “está chovendo DNA lá fora”, para se referir às sementes do salgueiro levadas pelo vento.

A pluma de celulose, apesar de volumosa, é simplesmente um paraquedas a ser descartado. Toda a exibição, com o algodão, os amentilhos, a árvore em si, tem apenas um propósito: espalhar o DNA pelo campo. Não é qualquer DNA, mas sim o DNA cujos caracteres codificados formam instruções específicas para a construção de salgueiros que gerarão uma nova geração de sementes felpudas. Essas pequenas partículas fofas estão, literalmente, disseminando instruções para se reproduzirem. Estão presentes porque seus ancestrais tiveram sucesso em fazer o mesmo. Está chovendo instruções lá fora; está chovendo programas; está chovendo crescimento de árvores, dispersão de algodão, algoritmos. Isso não é uma metáfora, é a pura verdade. Não poderia ser mais claro se estivesse chovendo disquetes. (DAWKINS, 1985. p. 111)

Os cérebros das entidades biológicas se desenvolveram para serem altamente sensíveis a padrões. Um mecanismo que algumas vezes leva a equívocos, é certo. Mas em geral tem permitido que a estratégia da inteligência prevaleça perante outras como a da velocidade, do tamanho, do veneno mais letal ou das garras ou dentes mais afiados na corrida evolutiva. O que não deve ser surpreendente, posto que proporciona meios mais eficientes de se atingir os

objetivos necessários para sobrevivência, como nota Richard Dawkins, em seu livro publicado originalmente em 1976, *The Selfish Gene* (2006):

Máquinas de sobrevivência que podem simular o futuro [em suas mentes] estão um passo à frente das máquinas de sobrevivência que só podem aprender com base em tentativa e erro. O problema com a tentativa é que leva tempo e energia. O problema com o erro é que frequentemente é fatal. A simulação é mais segura e mais rápida. (p.59). Tradução nossa.

De fato, um animal, digamos um pássaro, que reconhece padrões como a silhueta de uma fêmea será mais bem sucedido em promover seus genes para próxima geração. Mesmo que algumas vezes se confunda, por um momento, reconhecendo em uma folha ou uma rocha, sob certo ângulo, como se fosse o contorno de uma fêmea.

No livro *On Intelligence*, Jeff Hawkins(2004), o argumento central é que a capacidade de previsão é o elemento fundamental da cognição. Ele propõe que o cérebro humano funciona como um sistema de predição, constantemente fazendo estimativas e gerando modelos internos do mundo para antecipar eventos futuros. Essa abordagem difere da visão tradicional de que o cérebro é apenas um processador de informações estáticas. Hawkins destaca que a predição é essencial para a percepção, o planejamento e a tomada de decisões, permitindo que os seres humanos se adaptem e ajam de forma eficaz em seu ambiente.

Nesta linha teórica, o sistema nervoso extrai características do ambiente, seleciona aspectos relevantes, os analisa e oferece, para tomada de decisão, um número limitado de opções. E é necessário que assim o seja, a realidade, afinal, oferece uma enormidade de informações que nenhum poder computacional biológico ou artificial poderia dar conta em sua totalidade: ondas eletromagnéticas dos mais diversos comprimentos, moléculas das mais diversas composições arrastadas pelo ar, objetos que variam de tamanho desde dimensões microscópicas a enormidades espalhados pelo ambiente. Todas as decisões assim são tomadas com um limitado acesso ao real, mediado por órgãos sensoriais e pelas capacidades cognitivas.

De acordo com Solomonoff, o algoritmo ótimo para explicar um conjunto de dados é geralmente aquele que pode ser descrito de forma mais concisa e eficiente. Por sua vez, a MDL não apenas destaca a interdependência entre a descrição do modelo e o objeto que ele descreve, mas também busca identificar a descrição mais

concisa que incorpora os dados, eliminando elementos aleatórios e redundantes que não agregam informações significativas. Além disso, a KC fornece uma medida da aleatoriedade, permitindo identificar se um conjunto de dados ou programa apresenta padrões discerníveis ou é genuinamente aleatório. Em outras palavras, enquanto elementos aleatórios em um conjunto de dados podem ser considerados ruído e descartados por não conterem informações relevantes para revelar um padrão, a descrição mais concisa exonera quaisquer padrões.

A disciplina da inferência estatística também se depara com a tarefa essencial de descobrir padrões de uma amostra de dados. Muitos destes dados podem ser irrelevantes ou pouco relevantes para a explicação de um fenômeno que desejamos compreender. Uma das abordagens mais simples é utilizar o método de análise das componentes principais. Este procedimento permite organizar os dados em função da correlação entre eles, estabelecendo vetores ortogonais que capturam a maior parte da variância dentro de um espaço multidimensional. A primeira componente principal captura a maior parte da variância presente nos dados, enquanto as subsequentes capturam as variações remanescentes, de forma progressivamente menor. Assim, é possível identificar padrões, estruturas e tendências nos dados, bem como reduzir a complexidade do conjunto de dados, tornando-o mais fácil de visualizar e interpretar para, assim, se obter uma alta generalidade.

É coerente considerar que métodos com função similar agem moldando a evolução e, conseqüentemente, a cognição dos organismos biológicos. A sobrevivência dos organismos vivos requer uma adaptação contínua, eficiente e ágil a essas mudanças, visando garantir o sucesso reprodutivo, do qual decorre a continuidade de uma espécie no tempo.

Leslie Valiant (2013) cunha o termo *ecorritmos*, algoritmos baseados em princípios biológicos, que encontra uma interseção interessante com a AIT. Ecorritmos descrevem comportamentos adaptados de organismos biológicos na forma de uma *função* em relação uma grande quantidade de *variáveis*. Em oposição aos algoritmos desenvolvidos por um programador, o ecorritmo emerge naturalmente através do processo de seleção natural, no qual processos estocásticos possuem significativa influência. Em vez de se basearem em uma teoria prévia para resolver um determinado problema, os ecorritmos não possuem uma teoria

explícita, apenas precisam ser eficientes o suficiente para enfrentar os desafios que os organismos biológicos encontram.

A AIT permite uma investigação abrangente das estratégias de codificação da informação genética, a eficiência de sua transmissão ao longo das gerações e o processamento dessas informações durante o desenvolvimento ontogenético. Deste modo, a AIT se apresenta como um instrumento promissor para decifrar a informação contida nos algoritmos biológicos que orientam a sobrevivência, a adaptação e a capacidade cognitiva dos organismos.

Além disso, no contexto da ciência e da computação, destaca-se a importância de se considerar a complexidade e a imprevisibilidade na análise de sistemas de informação. Desta maneira, a AIT se mostra como uma ferramenta valiosa para enfrentar os desafios de um mundo cada vez mais interconectado, oferecendo novas maneiras de compreender e descrever a complexidade subjacente não apenas aos sistemas naturais, mas também aos novos sistemas artificiais, como internet das coisas, redes neurais artificiais, modelos de linguagem, entre outros.

Concluimos que este trabalho atingiu seus objetivos ao proporcionar uma visão abrangente do papel e do potencial da AIT em diversos contextos. Especificamente, destacamos a compreensão aprofundada dos limites do conhecimento humano, da natureza da realidade e da importância da complexidade e da aleatoriedade na compreensão científica do mundo. No entanto, é importante ressaltar que a exploração mais aprofundada das implicações dos teoremas da incompletude de Gödel e a revisão da extensa bibliografia coletada ao longo da pesquisa permanecem como oportunidades para futuras iterações e continuidade desta pesquisa.

REFERÊNCIAS

- AGAPOW, P.-M. "Building a better worm: Organisation and evolution in development", **Complexity International**, Cited by: 0, v. 12, p. 1 – 10, 2008. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-77949340763&partnerID=40&md5=56e64b7881a0561d571329046d5201c7>.
- ANÉ, C., SANDERSON, M. J. "Missing the forest for the trees: Phylogenetic compression and its implications for inferring complex evolutionary histories", **Systematic Biology**, Cited by: 39; All Open Access, Green Open Access, v. 54, n. 1, p. 146 – 157, 2005. DOI: 10.1080/10635150590905984. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-15944396106&doi=10.1080%2F10635150590905984&partnerID=40&md5=5cf638a7cd660341f58b3d17318ee808>.
- BACAËR, N., "Verhulst and the logistic equation (1838)". **A Short History of Mathematical Population Dynamics**, London, Springer London, 2011. p. 35–39. DOI: 10.1007/978-0-85729-115-8_6. Disponível em: https://doi.org/10.1007/978-0-85729-115-8_6.
- BASHARIN, G. P., LANGVILLE, A. N., NAUMOV, V. A. "The life and work of A.A. Markov", **Linear Algebra and its Applications**, v. 386, n. 1- 3 SUPPL., p. 3–26, jul. 2004. DOI: 10.1016/j.laa.2003.12.041. Disponível em: www.elsevier.com/locate/laa. Acesso em: 25 maio 2023.
- BATESON, G. **Mind and nature : a necessary unity**. 1. ed. New York, US, E. P. Dutton, 1979.
- BEALL, J., GLANZBERG, M., RIPLEY, D., "Liar Paradox". In: ZALTA, E. N. (Org.), **The Stanford Encyclopedia of Philosophy**, {F}all 202 ed. [S.l.], Metaphysics Research Lab, Stanford University, 2020. .
- BRODA, E. "The interaction of Boltzmann with Mach, Ostwald and Planck and his influence on Nerst and Einstein", **16th International Congress on the History of Science (Bucharest)**, 1981. .
- CHAITIN, G. J. "A Theory of Program Size Formally Identical to Information Theory", **Journal of the ACM (JACM)**, v. 22, n. 3, p. 329–340, 1 jul. 1975. DOI: 10.1145/321892.321894. .
- CHAITIN, G. J. "Information-Theoretic Limitations of Formal Systems", **Journal of the ACM (JACM)**, v. 21, n. 3, p. 403–424, 1974. DOI: 10.1145/321832.321839. .
- CILIBRASI, R. L., VITÁNYI, P. M. B. "Fast Phylogeny of SARS-CoV-2 by Compression", **Entropy**, v. 24, n. 4, p. 439, 2022. DOI: 10.3390/e24040439. .
- COROMINAS-MURTRA, B., SEOANE, L. F., SOLÉ, R. "Zipf's Law, unbounded complexity and open-ended evolution", **Journal of the Royal Society Interface**, v. 15, n. 149, p. 20180395, 2018. DOI: 10.1098/rsif.2018.0395. .
- DAVIS, M. **Engines of logic: mathematicians and the Origin of th Computer**. 2nd. ed. New York, W. W. Norton & Company, 2001.
- DAWKINS, R. **The Blind Watchmaker**. 1st. ed. New York, Norton, 1985.
- DAWKINS, R. **The Selfish Gene**. 30th Anive ed. [S.l.], Oxford University Press, 2006.
- DĘBOWSKI, Ł. "Is natural language a perigraphic process? The theorem about facts and words revisited", **Entropy**, Cited by: 14; All Open Access, Gold Open Access, Green Open Access, v. 20, n. 2, 2018. DOI: 10.3390/e20020085. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85053700164&doi=10.3390%2Fe20020085&partnerID=40&md5=00e1e20e43c9127d0762e884f46441b6>.
- DEVINE, S. D. "Understanding how replication processes can maintain systems away from equilibrium using Algorithmic Information Theory", **BioSystems**, v. 140, p. 8–22, 2016. DOI: 10.1016/j.biosystems.2015.11.008. .
- DOTY, D. "Every sequence is decompressible from a random one", **Lecture Notes in Computer**

Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), v. 3988 LNCS, n. November, p. 153–162, 2006. DOI: 10.1007/11780342_17. .

DYSON, G. **Turing's Cathedral: The Origins of the Digital Universe**. 1st. ed. New York, Vintage Books, 2012.

GAUVRIT, N., SOLER-TOSCANO, F., GUIDA, A. "A preference for some types of complexity comment on "perceived beauty of random texture patterns: A preference for complexity"", **Acta Psychologica**, Cited by: 7, v. 174, p. 48–53, 2017. DOI: 10.1016/j.actpsy.2017.01.007. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85011664718&doi=10.1016%2Fj.actpsy.2017.01.007&partnerID=40&md5=96c3b74064a8c211615b8a2acb610986>.

GEFTER, A. **The Man Who Tried to Redeem the World with Logic - Nautilus**. 29 jan. 2015. Nautilus:Technology. Disponível em: <https://nautilus.us/the-man-who-tried-to-redeem-the-world-with-logic-235253/>. Acesso em: 23 maio 2023.

GLEICK, J. **Chaos: Making a new science**. New York, Viking, 1987.

GOODWIN, R. M., "A Growth Cycle". **Essays in Economic Dynamics**, London, Palgrave Macmillan UK, 1982. p. 165–170. DOI: 10.1007/978-1-349-05504-3_12. Disponível em: https://doi.org/10.1007/978-1-349-05504-3_12.

HAND, L. N., FINCH, J. D. **Analytical mechanics**. 1. ed. Cambridge UK, Cambridge University Press, 1998.

HAWKINS, J., BLAKESLEE, S. "On intelligence", **Neural Networks**, p. 272, 2004. Disponível em: www.onintelligence.com.

HJORTH-JENSEN, M. "COMPUTATIONAL PHYSICS Morten Hjorth-Jensen", **History**, 2009. .

HOFSTADTER, D. R. **Gödel, Escher, Bach: An Eternal Golden Braid**. 20th Year ed. New York, NY, Basic Books, 1999.

HOUZÉ, É., DESSALLES, J. L., DIACONESCU, A., *et al.* "What Should I Notice? Using Algorithmic Information Theory to Evaluate the Memorability of Events in Smart Homes", **Entropy**, v. 24, n. 3, p. 346, 2022. DOI: 10.3390/e24030346. .

HUTTER, M. "A gentle introduction to the universal algorithmic agent AIXI", **Artificial General Intelligence**, n. January, 2003. Disponível em: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.8797%5Cnhttp://www.researchgate.net/publication/228851452_A_gentle_introduction_to_the_universal_algorithmic_agent_AIXI/file/9fcfd513fb2bb4148d.pdf.

HUTTER, M. "Open problems in universal induction & intelligence", **Algorithms**, v. 2, n. 3, p. 879–906, 2009. DOI: 10.3390/a2030879. .

IRVINE, A. D., DEUTSCH, H., "Russell's Paradox". In: ZALTA, E. N. (Org.), **The Stanford Encyclopedia of Philosophy**, {S}pring 2 ed. [S.l.], Metaphysics Research Lab, Stanford University, 2021. .

JOHNSTON, I. G., DINGLE, K., GREENBURY, S. F., *et al.* "Symmetry and simplicity spontaneously emerge from the algorithmic nature of evolution", **Proceedings of the National Academy of Sciences of the United States of America**, v. 119, n. 11, p. e2113883119--e2113883119, 2022. DOI: 10.1073/pnas.2113883119. .

KANG, C. G. "Origin of stability analysis in view of On Governors by maxwell", **Journal of Institute of Control, Robotics and Systems**, v. 22, n. 6, p. 435–444, 2016. DOI: 10.5302/J.ICROS.2016.16.0057. .

KLEENE, S. C., "Turing's Analysis of Computability and Major Application of it". In: HERKEN, R. (Org.), **The Universal Turing Machine: a half-century Survey**, [s.l.], Oxford University Press, 1988. p. 17–54.

KOVALSKI, E. L. **Diálogos entre Wetware e Inteligência Artificial: Implicações das concepções sobre processos cognitivos biológicos no desenvolvimento da tecnologia computacional . Rio de Janeiro**. 2021. 119 f. UFRJ, 2021. Disponível em:

http://146.164.248.81/hcte/docs/dissertacoes/2021/erickson_leon_kovalski.pdf.

LECUN, Y., BENGIO, Y., HINTON, G. "Deep learning", **Nature**, v. 521, n. 7553, 2015. DOI: 10.1038/nature14539. .

LETTVIN, J. Y., MATURANA, H. R., MCCULLOCH, W. S., *et al.* "What the Frog's Eye Tells the Frog's Brain. Proceedings of the", **Proceedings of the IRE**, v. 47, n. 11, p. 1940–1959, 1959. Disponível em: <https://ieeexplore.ieee.org/document/4065609>.

LORENZ, E. N. "Predictability: does the flap of a butterfly's wings in Brazil set off a tornado in Texas?". 1972. **Anais [...]** [S.l: s.n.], 1972. p. 1–3. Disponível em: <https://fermatlibrary.com/s/predictability-does-the-flap-of-a-butterflys-wings-in-brazil-set-off-a-tornado-in-texas>. Acesso em: 8 jun. 2023.

MAGUIRE, P., MOSER, P., MAGUIRE, R. "Understanding consciousness as data compression", **Journal of Cognitive Science**, Cited by: 13; All Open Access, Gold Open Access, Green Open Access, v. 17, n. 1, p. 63–94, 2016. DOI: 10.17791/jcs.2016.17.1.63. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85009959357&doi=10.17791%2Fjcs.2016.17.1.63&partnerID=40&md5=ed9d6cfdcefb1c8ce926643f7155302>.

MARKS, R. J., CHAITIN, G. J. **How did Ray Solomonoff Kickstart Algorithmic Information Theory?** 15 mar. 2021. Mindmatters.ai. Disponível em: <https://mindmatters.ai/2021/03/how-did-ray-solomonoff-kickstart-algorithmic-information-theory/>. Acesso em: 11 jun. 2023.

MCCULLOCH, W. S., PITTS, W. "A logical calculus of the ideas immanent in nervous activity", **The Bulletin of Mathematical Biophysics**, v. 5, n. 4, p. 115–133, 1943. DOI: 10.1007/BF02478259. .

MIŁKOWSKI, M. "Is evolution algorithmic?", **Minds and Machines**, v. 19, n. 4, p. 465–475, 2009. DOI: 10.1007/s11023-009-9170-6. .

MITCHELL, M. **Complexity: A Guided Tour**. 1 (Kindle) ed. Oxford, Oxford University Press, 2009.

NEUMANN, J. von. "First draft report on EDVAC", **IEEE Ann. Hist. Comput.**, v. 15, n. 4, p. 27–75, 1993. .

OEXLE, K. "Data compressibility, physical entropy, and evolutionary a priori relation between observer and object", **Physical Review E**, Cited by: 1, v. 51, n. 3, p. 2651 – 2653, 1995. DOI: 10.1103/PhysRevE.51.2651. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-35949006914&doi=10.1103%2FPhysRevE.51.2651&partnerID=40&md5=993dad7719a06ee6bca5d79136c6ce6f>.

PINGUELLI, L. R. **Tecnociências e humanidades: Novos Paradigmas, Velhas Questões**. 1. ed. Rio de Janeiro, Paz e Terra, 2008.

POLAND, J., HUTTER, M. "Asymptotics of discrete MDL for online prediction", **IEEE Transactions on Information Theory**, Cited by: 17; All Open Access, Green Open Access, v. 51, n. 11, p. 3780 – 3795, 2005. DOI: 10.1109/TIT.2005.856956. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-27744462709&doi=10.1109%2FTIT.2005.856956&partnerID=40&md5=3f20451cad58614332b8722d492945bc>.

PRIGOGINE, I. **O Nascimento do Tempo**. 1. ed. São Paulo, Edições 70, 1988.

RISSANEN, J. "Modeling by shortest data description", **Automatica**, v. 14, n. 5, p. 465–471, 1978. DOI: 10.1016/0005-1098(78)90005-5. .

ROSENBLATT, F. "The Design of an Intelligent Automaton", **U.S. Office of Naval Research**, v. 6, n. 2, p. 7, 1958. .

ROSENBLUTH, A., WIENER, N., BIGELOW, J. **Behaviour, purpose and teleology. Philosophy of Science**. [S.l: s.n.], 1943

SCHEFFEL, J. "On the Solvability of the Mind–Body Problem", **Axiomathes**, Cited by: 1; All Open Access, Green Open Access, Hybrid Gold Open Access, v. 30, n. 3, p. 289–312, 2020. DOI: 10.1007/s10516-019-09454-x. Disponível em:

<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85084924994&doi=10.1007%2Fs10516-019-09454-x&partnerID=40&md5=6919a8ac9f08aa3b6498548d6fb459c9>.

SHANNON, C. E. "A Mathematical Theory of Communication", **Bell System Technical Journal**, v. 27, n. 3, p. 379–423, jul. 1948. DOI: 10.1002/j.1538-7305.1948.tb01338.x. Disponível em: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6773024>.

SHEN, A. "Discussion on Kolmogorov complexity and statistical analysis", **Computer Journal**, Cited by: 15, v. 42, n. 4, p. 340 – 342, 1999. DOI: 10.1093/comjnl/42.4.340. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-0032685680&doi=10.1093%2Fcomjnl%2F42.4.340&partnerID=40&md5=3c0cbd97a5769a102e3f01cf3d369e01>.

SIGMUND, K., HOFSTADTER, D. R. **Exact Thinking in Demented Times: The Vienna Circle and the Epic Quest for the Foundations of Science**. 1. ed. New York, Hachette Books, 2017.

SOLOMONOFF, R. J. "A formal theory of inductive inference. Part I", **Information and Control**, v. 7, n. 1, p. 1–22, 1964a. DOI: 10.1016/S0019-9958(64)90223-2. .

SOLOMONOFF, R. J. "A formal theory of inductive inference. Part II", **Information and Control**, v. 7, n. 2, p. 224–254, 1964b. .

SOLOMONOFF, R. J. "A Preliminary Report On a General Theory of Inductive Inference", n. 138, 1960. .

STEBBING, A. R. D. "Genetic parsimony: A factor in the evolution of complexity, order and emergence", **Biological Journal of the Linnean Society**, v. 88, n. 2, p. 295–308, 2006. DOI: 10.1111/j.1095-8312.2006.00652.x. .

STERKENBURG, T. F. "Solomonoff prediction and Occam's razor", **Philosophy of Science**, Cited by: 7; All Open Access, Green Open Access, v. 83, n. 4, p. 459–479, 2016. DOI: 10.1086/687257. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84987677428&doi=10.1086%2F687257&partnerID=40&md5=41976d169b39db031e85ed99a5957d39>.

STONE, N. C., LEIGH, N. W. C. "A statistical solution to the chaotic, non-hierarchical three-body problem", **Nature**, v. 576, n. 7787, p. 406–410, 2019. DOI: 10.1038/s41586-019-1833-8. Disponível em: <http://dx.doi.org/10.1038/s41586-019-1833-8>.

THRUN, S., BAXTER, J., THRUN, S. "A Bayesian/information theoretic model of learning to learn via multiple task sampling", **Machine Learning**, v. 28, p. 7–39, 1997. DOI: 10.1023/A:1007327622663. Disponível em: <http://link.springer.com/article/10.1023/A:1007327622663>.

TURING, A. M. "On computable numbers, with an application to the Entscheidungsproblem", **Proceedings of the London Mathematical Society**, v. 42, n. 2, p. 230–265, 1936. Disponível em: <http://plms.oxfordjournals.org/cgi/reprint/s2-42/1/230>.

VALIANT, L. **Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World**. 1st. ed. New York, Basic Books, 2013.

VERHEYEN, P. "From Information and Quantum Physics to Consciousness and Reality", **Sci**, v. 3, n. 4, p. 35, 2021. DOI: 10.3390/sci3040035. .

VITANYI, P. M. B. "Ray Solomonoff, Founding Father of Algorithmic Information Theory", **Algorithms**, v. 3, n. 3, p. 260–264, 2010. DOI: 10.3390/a3030260. .

WIENER, N. **Cybernetics: or the Control and Communication in the Animal and the Machine**. 2nd. ed. Cambridge, Massachusetts, MIT Press, 1961.

APÊNDICE A: CÓDIGO FONTE DOS EXEMPLOS E VISUALIZAÇÕES UTILIZADOS NESTE TRABALHO

Os códigos aqui disponíveis foram elaborados como parte de estudos e trabalhos acadêmicos realizados no decorrer do curso de Matemática e convertidos para linguagem Python com o auxílio da ferramenta GitHub Copilot.

Código em Python para gerar as visualizações do problema de três corpos:

```
import numpy as np
from scipy.integrate import solve_ivp
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D

# Masses
mass_sun = 1 # sun mass (in solar masses)
#mass_earth = 3.003e-6 # earth mass (in solar masses)
#mass_jupiter = 0.00095 # jupiter mass (in solar masses)
mass_earth = 1
mass_jupiter = 1 # three suns is the best for this visualization
masses = np.array([mass_sun, mass_earth, mass_jupiter])

# Initial Conditions
y0 = [-0.5000001, 0, 0, # x1, y1, z1
      0.5, 0, 0, # x2, y2, z2
      0, 0, 0, # x3, y3, z3
      0, .497, 0.2, # vx1, vy1, vz1 (Velocities are here)
      0, -1, 0.1, # vx2, vy2, vz2
      0, 0, 0.05] # vx3, vy3, vz3

# Definition of the Function
def Planets(t, y, masses):
    f = np.zeros(18)
    f[0:3] = y[9:12]
    f[3:6] = y[12:15]
    f[6:9] = y[15:18]

    for i in range(3):
        for j in range(3):
            if i != j:
                rij = np.sqrt((y[j*3]-y[i*3])**2 + (y[j*3+1]-
y[i*3+1])**2 + (y[j*3+2]-y[i*3+2])**2)
                f[9+i*3:9+i*3+3] += -masses[j]*(y[i*3:i*3+3] -
y[j*3:j*3+3]) / rij**3

    return f

N = 3000 # N are the Steps and T is the duration of each step...
means how much detail you get
T = 0.001 # N*T gives the total time of the simulation

t = np.linspace(0, N*T, N)
```

```

solution = solve_ivp(Planets, [0, N*T], y0, t_eval=t, rtol=1e-15,
args=(masses,))

# 3D Visualization
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')

ax.plot(solution.y[0], solution.y[1], solution.y[2], '-g') #
Positions body 1
ax.plot(solution.y[3], solution.y[4], solution.y[5], '-r') #
Positions body 2
ax.plot(solution.y[6], solution.y[7], solution.y[8], '-b') #
Positions body 3

ax.scatter(*y0[:3], color='g', label='body 1 start')
ax.scatter(*y0[3:6], color='r', label='body 2 start')
ax.scatter(*y0[6:9], color='b', label='body 3 start')

ax.set_xlabel('X')
ax.set_ylabel('Y')
ax.set_zlabel('Z')

#plt.legend() (labels are overrated)
plt.show()

```

Código em Python para gerar as visualizações da equação logística:

```

import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt

# Logistic function
def logistic(P, t, r, K):
    return r * P * (1 - P / K)

# Parameters
r = 0.1 # growth rate
K = 1000 # carrying capacity
P0 = 10 # initial population
t = np.linspace(0, 110, 1000) # time grid

# Solve ODE
P = odeint(logistic, P0, t, args=(r, K))

# Plot
plt.figure(figsize=(8, 6))
plt.plot(t, P, label="Logistic model")
plt.xlabel("Time")
plt.ylabel("Population")
plt.title("Logistic growth model: r={}, K={}".format(r, K))
plt.legend()
plt.grid(True)
plt.show()

```

Código em Python para gerar as visualizações do mapa logístico:

```
import numpy as np
import matplotlib.pyplot as plt

def logistic_map(x, r):
    return r * x * (1 - x)

# Initial condition
x_init = .43

# Values of r
r_values = [.9, 2.6, 3, 3.7, 3.9, 4] # Adding an r value close to the
onset of chaos

# Number of iterations
n_iterations = 150

for r in r_values:
    x = np.zeros(n_iterations)
    x[0] = x_init

    # Iterate the logistic map
    for i in range(1, n_iterations):
        x[i] = logistic_map(x[i-1], r)
        if np.isclose(x[i], 1):
            break

    # Plot
    plt.figure()
    plt.plot(range(i+1), x[:i+1], label=f'r={r}')
    plt.xlabel('Iteration')
    plt.ylabel('x')
    plt.legend()
    plt.title('Evolution of the logistic map')
    plt.show()
```

Código em Python para gerar as visualizações do diagrama de bifurcações:

```
import numpy as np
import matplotlib.pyplot as plt

def logistic_map(x, r):
    return r * x * (1 - x)

# Initial condition
x_init = 0.43

# Values of r
r_values = np.linspace(0.9, 4.0, 2500)

# Number of iterations and number of iterations to skip
n_iterations = 2500
n_skip = 250 # We'll skip the first n_skip iterations to allow the
system to reach equilibrium
```

```

last_values = []

for r in r_values:
    x = x_init

    # Iterate the logistic map
    for i in range(n_iterations):
        x = logistic_map(x, r)
        if i >= n_skip:
            last_values.append((r, x))

# Convert to numpy array and plot
last_values = np.array(last_values)
plt.figure(figsize=(40, 28))
plt.scatter(last_values[:, 0], last_values[:, 1], s=0.005)
plt.xlabel('r', fontsize=30)
plt.ylabel('Equilibrium population', fontsize=30)
plt.title('Bifurcation diagram of the logistic map', fontsize=35)
plt.tick_params(axis='both', which='major', labelsize=20)
plt.show()

```

Sistema Predador-Presa de Lotka-Volterra

```

import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt

# function that returns dy/dt
def model(z,t):
    x = z[0]
    y = z[1]
    dxdt = alpha*x - beta*x*y
    dydt = delta*x*y - gamma*y
    dzdt = [dxdt,dydt]
    return dzdt

# parameters
alpha = 1.1
beta = 0.4
gamma = 0.4
delta = 0.1

# initial condition
z0 = [10,5]

# time points
t = np.linspace(0,50,500)

# solve ODE
z = odeint(model,z0,t)

x = z[:,0]
y = z[:,1]

# plot results
plt.figure(figsize=(12,6))
plt.plot(t,x,'r-',label='Prey')
plt.plot(t,y,'b-',label='Predator')

```

```

plt.ylabel('Population')
plt.xlabel('Time')
plt.legend(loc='best')
plt.grid()
plt.show()

```

Neurônio de McCulloch-Pitts

```

import matplotlib.pyplot as plt
import numpy as np

class McCullochPittsNeuron:
    def __init__(self, weights, threshold):
        self.weights = weights
        self.threshold = threshold

    def activate(self, inputs):
        if len(inputs) != len(self.weights):
            raise ValueError("Input size must match the number of
weights")

        weighted_sum = sum(w * x for w, x in zip(self.weights,
inputs))

        if weighted_sum >= self.threshold:
            return 1
        else:
            return 0

class TripleLayerNetwork:
    def __init__(self):
        # Set weights for the neurons in the first and second layers
        self.first_layer = [
            McCullochPittsNeuron([1, 1], 2) # Neuron 1: weights [1,
1], threshold 2 for AND operation
        ]
        self.second_layer = [
            McCullochPittsNeuron([1, 1], 1) # Neuron 1: weights [1,
1], threshold 1 for OR operation
        ]
        self.third_layer = McCullochPittsNeuron([-1, -1], -1) #
Neuron 1: weights [-1, -1], threshold -1 for NAND operation

    def activate(self, inputs):
        if len(inputs) != 4:
            raise ValueError("Input size must be 4")

        # Perform the AND operation on the first two inputs
        and_output = self.first_layer[0].activate(inputs[:2])

        # Perform the OR operation on the last two inputs
        or_output = self.second_layer[0].activate(inputs[2:])

        # Combine the outputs of the AND and OR operations and perform
the NAND operation
        nand_input = [and_output, or_output]
        nand_output = self.third_layer.activate(nand_input)

```

```

        return nand_output

# Example usage
network = TripleLayerNetwork()

# Generate all possible inputs
inputs = np.array([[0, 0, 0, 0],
                  [0, 0, 0, 1],
                  [0, 0, 1, 0],
                  [0, 0, 1, 1],
                  [0, 1, 0, 0],
                  [0, 1, 0, 1],
                  [0, 1, 1, 0],
                  [0, 1, 1, 1],
                  [1, 0, 0, 0],
                  [1, 0, 0, 1],
                  [1, 0, 1, 0],
                  [1, 0, 1, 1],
                  [1, 1, 0, 0],
                  [1, 1, 0, 1],
                  [1, 1, 1, 0],
                  [1, 1, 1, 1]])

# Activate the network with all possible inputs
outputs = np.array([network.activate(input) for input in inputs])

# Visualize the results
plt.figure(figsize=(16, 6))
plt.scatter(range(len(outputs)), outputs, s=1000) # Increase plot
point size
plt.xticks(range(len(inputs)), [str(input) for input in inputs])
plt.xlabel('Inputs', fontsize=16)
plt.ylabel('Output', fontsize=16)
plt.title('Example of McCulloch-Pitts Logic Calculus', fontsize=18)
plt.ylim([-0.2, 1.2])
plt.yticks([0, 1]) # Set y-axis ticks to only 0 and 1
plt.grid(True)
plt.show()

```

Random Walk

```

import numpy as np
import matplotlib.pyplot as plt

# Define the transition probabilities as a 3x3 matrix
transition_matrix = np.array([[0.2, 0.5, 0.3],
                              [0.9, 0.05, 0.05],
                              [0.4, 0.3, 0.3]])

# Define the labels for the nodes
node_labels = ['A', 'B', 'C']

# Set the initial state
current_state = 0 # Starting from node A

# Perform the random walk simulation
num_steps = 10000 # Number of steps in the random walk

```

```

visit_counts = [0, 0, 0] # Initialize the visit counts for each node
probabilities = [] # Store the probabilities at each 10 steps

print("Random Walk Simulation:")
print("Starting from Node A")

for step in range(1, num_steps + 1):
    # Perform a random transition based on the transition
    probabilities
    current_state = np.random.choice(range(len(node_labels)),
    p=transition_matrix[current_state])

    # Increment the visit count for the current state
    visit_counts[current_state] += 1

    # Calculate the probabilities at each 10 steps
    if step % 10 == 0:
        total_steps = sum(visit_counts)
        step_probabilities = [count / total_steps for count in
visit_counts]
        probabilities.append(step_probabilities)

# Convert the probabilities list into a NumPy array
probabilities = np.array(probabilities)

# Plot the probabilities for each node at each 10 steps
x = np.arange(10, num_steps + 1, 10)
for i, node in enumerate(node_labels):
    plt.plot(x, probabilities[:, i], label=f'Node {node}')

# Set the plot title and labels
plt.title('Random Walk')
plt.xlabel('Number of Steps')
plt.ylabel('Probability')
plt.legend()

# Display the plot
plt.show()

```

Turing Machine

```

def turing_increment(binary_number):
    rules = {
        'Start': {'1': ('0', 'Right', 'Carry'), '0': ('1', 'N/A',
'Halt')}, 'B': {'1': ('N/A', 'Halt')},
        'Carry': {'1': ('0', 'Right', 'Carry'), '0': ('1', 'N/A',
'Halt')}, 'B': {'1': ('N/A', 'Halt')}}
    }

    print("Input Value:", binary_number)
    print("Input Value (Decimal):", int(binary_number,2))

    print("\n\nTuring Machine Rules:")
    print("{:<10} {:<10} {:<10} {:<10} {:<10}".format('Current State
|', 'Symbol Read |', 'Symbol Write |', 'Move Direction |', 'Next
State'))
    for state, transitions in rules.items():
        for symbol, action in transitions.items():

```

```

        print("{:<20} {:<15} {:<10} {:<15} {:<20}".format(state,
symbol, *action))
    print()

    # Reverse the list and append a blank at the end.
    tape = list(binary_number)[::-1] + ['B']
    head_position = 0
    state = 'Start'

def print_tape():
    for i, cell in enumerate(tape):
        if i == head_position:
            print(f"*{cell}", end="")
        else:
            print(f"|{cell}", end="")
    print("|")

while state != 'Halt':
    symbol = tape[head_position]

    print('State:', state)
    print('Symbol:', symbol)
    print_tape()
    print()

    symbol_write, _, state_next = rules[state][symbol]
    tape[head_position] = symbol_write

    if symbol == '1' and state == 'Start':
        state = 'Carry'
    else:
        state = state_next

    if symbol_write == '0':
        head_position += 1

    # Print the final tape, state, and head position.
    print('Final State:', state)
    print_tape()

    final_output_binary = ''.join(tape)[::-1].strip('B')
    print("\n\nOutput Value (binary):", final_output_binary)
    print("Output Value (decimal):", int(final_output_binary, 2))

# Test the Turing machine with a binary number.
k=input('Enter a binary number: ')
turing_increment(k)

```

APÊNDICE B: PUBLICAÇÕES EM TEORIA ALGORÍTMICA DA INFORMAÇÃO

Realizamos este estudo com o objetivo de proporcionar uma análise quantitativa da produção científica relacionada à Teoria da Informação Algorítmica (AIT). Utilizando a base de dados Scopus, selecionamos uma amostra de 207 artigos publicados em língua inglesa no período de 1973 a 2022. Através da contagem de palavras nos títulos e resumos, identificamos os 50 termos mais significativos, que servem como indicadores dos temas de pesquisa mais representativos nessa área.

Além disso, realizamos uma análise das áreas de pesquisa utilizando as informações fornecidas pela Plataforma Scopus. Essa análise revelou que a maioria dos artigos está relacionada às áreas de Matemática e Ciência da Computação. No entanto, também observamos a presença de pesquisas em outras áreas, como Artes e Humanidades e Psicologia. Para visualizar a distribuição das áreas, criamos gráficos que evidenciam a predominância da Ciência da Computação e da Matemática.

Por meio da aplicação de técnicas de regressão linear, também examinamos a evolução temporal da produção científica em AIT. Os resultados indicaram um aumento no número de publicações ao longo do tempo, demonstrando o crescente interesse e relevância dessa área multidisciplinar.

Essas informações sobre os principais temas abordados e a distribuição das áreas de pesquisa destacam a importância da AIT como um campo de estudo em constante expansão. Os resultados contribuem para uma compreensão mais abrangente da influência da AIT em diferentes campos de conhecimento e fornecem elementos para futuras investigações nessa área multidisciplinar.

Metodologia da coleta e tratamento de dados

Utilizando a base de dados Scopus, da editora Elsevier, realizamos a pesquisa pelos artigos que explicitamente incluem no título ou resumo as palavras “Algorithmic Information Theory”. O limite inferior da base é o ano de 1973 e estabelecemos o limite superior o ano de 2022, último ano completo durante a elaboração desta pesquisa (Quadro 1).

Quadro 1: Argumento de pesquisa na Plataforma Scopus (Elsevier)

```
Your query : TITLE-ABS-KEY ( "Algorithmic Information Theory" ) AND  
PUBYEAR > 1972 AND PUBYEAR < 2023 AND ( LIMIT-TO ( DOCTYPE , "ar" ) ) AND  
( LIMIT-TO ( LANGUAGE , "English" ) )
```

Foram selecionados apenas resultados referentes a artigos e publicados em língua inglesa, obtendo uma amostra de 207 artigos (lista completa no final do texto). A seguir os resultados foram exportados para o *software* gerenciador de referências Zotero versão 6. Utilizando funções do software produzimos uma listagem dos resumos que exportamos formato de texto para realizar a contagem de palavras. Para tanto utilizamos a biblioteca python *collections*. A seguir, manualmente analisamos os resultados excluindo as palavras de uso comum mantendo apenas as mais significativas que aparecem ao menos 10 vezes, que, esperamos, aparece em aproximadamente 5% dos artigos. Reconhecemos que a frequência real pode ser menor, pois artigos que tratam de um determinado tema tendem a repetir as palavras mais relevantes. No entanto, o objetivo foi identificar os termos mais significativos que representam os temas de pesquisa mais relevantes na área. As palavras "Teoria, "Informação" e "Algoritmo", assim como seus plurais e palavras similares foram excluídas por serem o termo de pesquisa inicial (isto é, aparecem ao menos uma vez em cada artigo). Colocamos os termos em ordem alfabética e consolidamos termos lexicograficamente similares do mesmo conceito. Selecionamos então os 50 termos mais significativos.

Em termos de distribuição de área de pesquisa, utilizando as ferramentas de análise de dados da própria plataforma Scopus. A seguir, extraímos os dados da plataforma Scopus e produzimos um gráfico utilizando software Excel, representando as áreas mais comuns e agrupando as áreas com menos de 5% de representatividade.

Finalmente organizamos as publicações por data e através de regressão linear, disponível no próprio software Excel, ajustamos uma linha de tendência.

Análise da amostra

A amostra de 207 artigos selecionados proporcionou uma visão abrangente dos temas e áreas de pesquisa mais representativos nesse campo de estudo. Ao examinar os títulos e resumos dos artigos, identificamos os 50 termos mais significativos (Tabela 2) que servem como indicadores dos temas mais abordados na AIT. Esses termos mostram conexão a temas centrais da AIT como compressão de dados, complexidade algorítmica, teoria da informação, algoritmos universais, entre outros. Essa seleção de termos fornece uma visão geral dos principais tópicos explorados pelos pesquisadores no campo da AIT (Gráfico 1 e Figura 1).

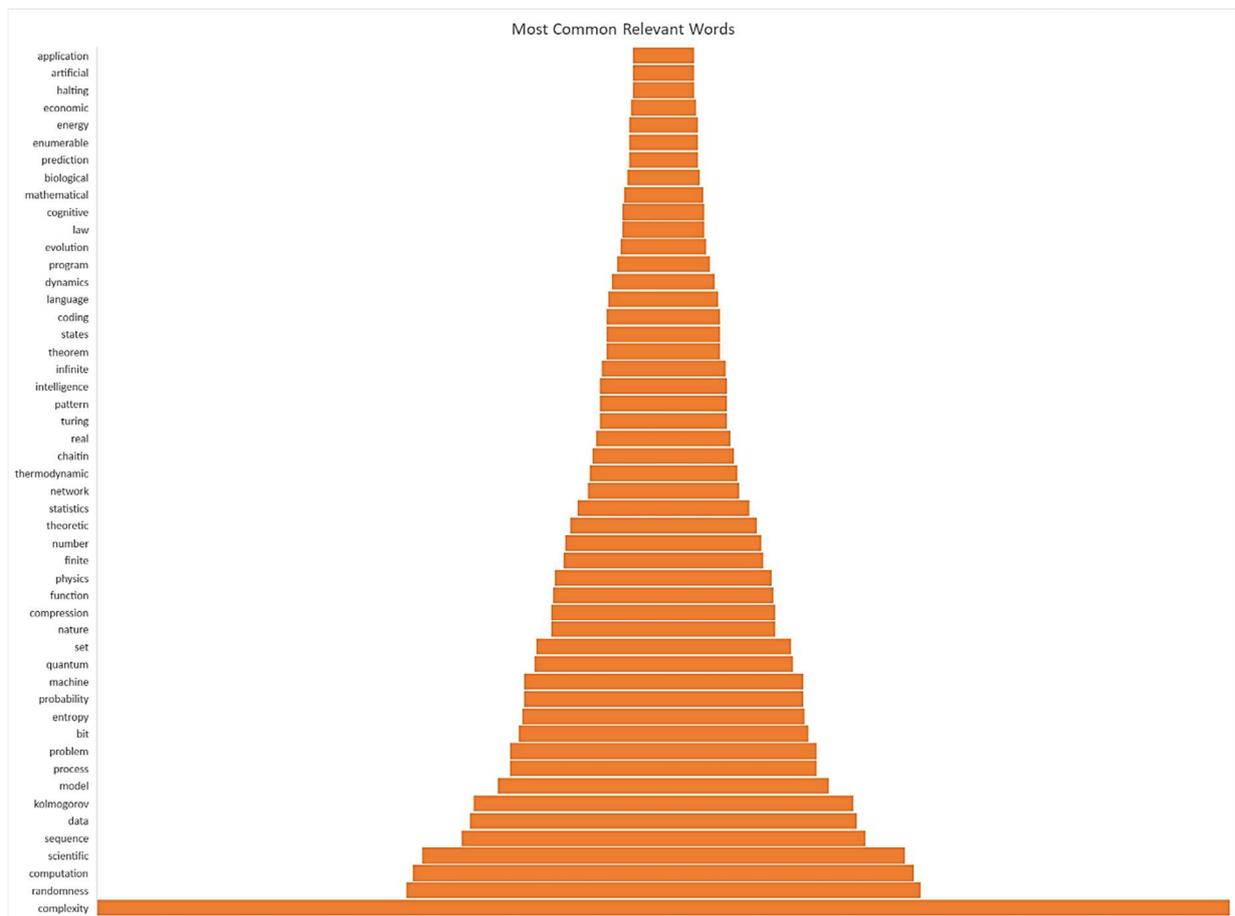


Gráfico1: Distribuição das palavras significativas mais comuns entre o resumo de 207 artigos publicados desde 1973, disponíveis na base de dados Scopus (Elsevier), excluídos "Teoria", "Informação" e "Algoritmo". Complexidade aparece em 10% dos resultados. Elaboração do autor.

Quanto à distribuição das áreas de pesquisa, observamos que a maioria dos artigos está relacionada às áreas de Matemática e Ciência da Computação, o que era esperado, considerando a natureza algorítmica e teórica da AIT. No entanto, também identificamos uma diversidade de áreas adicionais, incluindo Artes e Humanidades, Psicologia e outras disciplinas (Gráfico 2). É importante destacar que um mesmo artigo pode ser classificado em mais de uma área, o que ressalta a natureza interdisciplinar da AIT e sua aplicabilidade em diversos campos do conhecimento. Cerca de 55% dos artigos estão associados à Ciência da Computação e 50% à área da Matemática. Física e Astronomia aparece em quase 20% dos artigos e engenharia 12%. A seguir, 10% dos artigos se referem às áreas de Artes e Humanidades e quase 9% às Ciências Sociais. Aproximadamente 7% dos artigos são associados a áreas multidisciplinares. As áreas menos representadas podem ser vistas na Tabela 1.

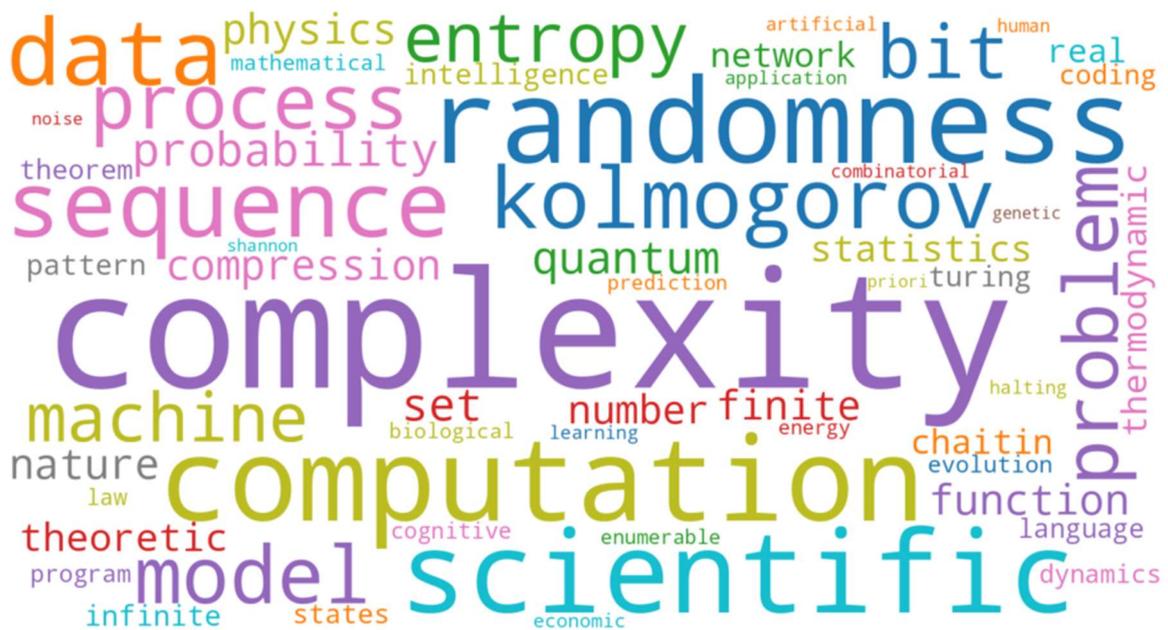


Figura 1: Palavras relevantes mais comuns em resumos de 207 artigos publicados desde 1973, disponíveis na base de dados Scopus (Elsevier), excluídos "Teoria", "Informação" e "Algoritmo". O tamanho da fonte na palavra reflete a frequência encontrada. Elaboração do autor.

PUBLICATIONS ON ALGORITHMIC INFORMATION THEORY (1973 - 2022)

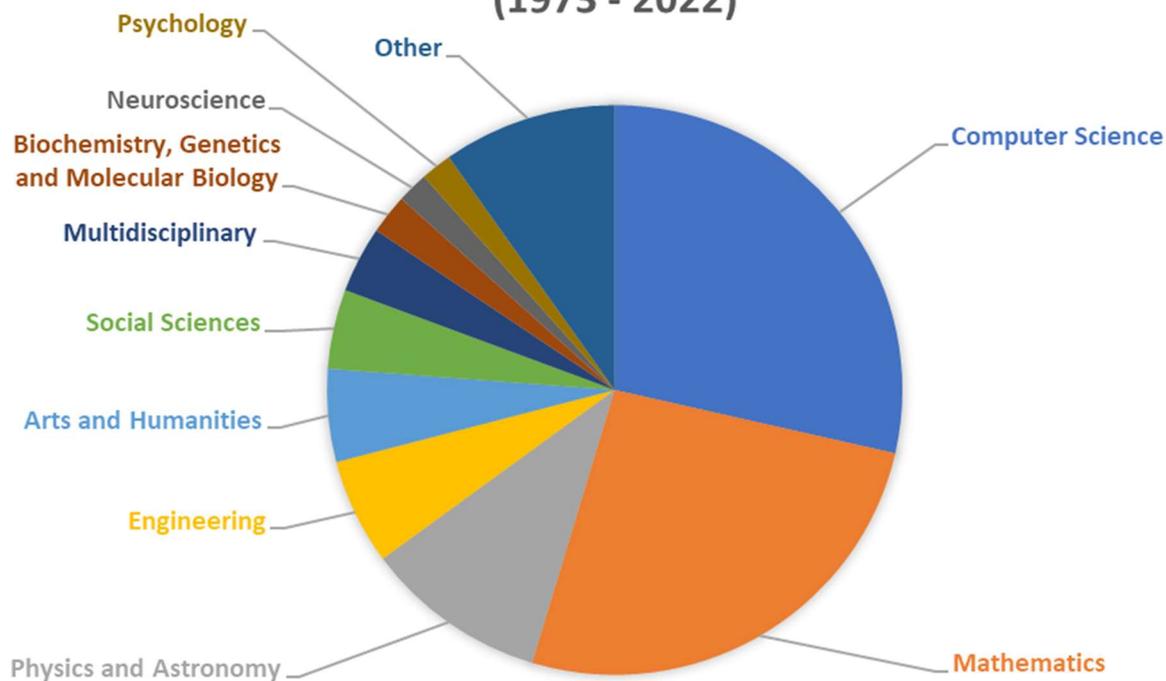


Gráfico 2: Visualização das principais áreas que utilizam a AIT em nossa amostra.

Tabela 1: Áreas com menor representatividade da AIT em nossa amostra segundo categorias da plataforma Scopus.

Área	Percentual de artigos relacionados
Bioquímica, genética e Biologia Molecular	4,4%
Neurociências	3,4%
Psicologia	3,4%
Agricultura e Ciências Biológicas	19%
Ciências da decisão	
Medicina	
Economia, Econometria e Finanças	
Química	
Ciências da Terra e Planetárias	
Ciências do Meio Ambiente	
Profissões da Saúde	

Ao analisar a evolução temporal da produção científica, utilizamos uma regressão linear para traçar uma linha de tendência, a qual apresenta coeficiente angular positivo para o período ($\beta=0,22$). Desta podemos depreender que apesar da variação, ano após ano, existe um visível crescimento em interesse pelo tema em nossa amostra.

A identificação dos termos mais relevantes destaca os principais tópicos de pesquisa e áreas de interesse na AIT, auxiliando pesquisadores a direcionarem seus estudos. Além disso, a diversidade de áreas em que a AIT é aplicada ressalta sua relevância e impacto em múltiplos domínios de conhecimento. A análise temporal, por sua vez, indica um crescente reconhecimento da importância da AIT como uma ferramenta teórica para a compreensão de problemas no âmbito da complexidade (Gráfico 3). Essa tendência sugere um futuro promissor para a pesquisa em AIT, com potenciais avanços e contribuições para diversas áreas acadêmicas e setores da indústria ligada a compressão, transmissão e encriptação.

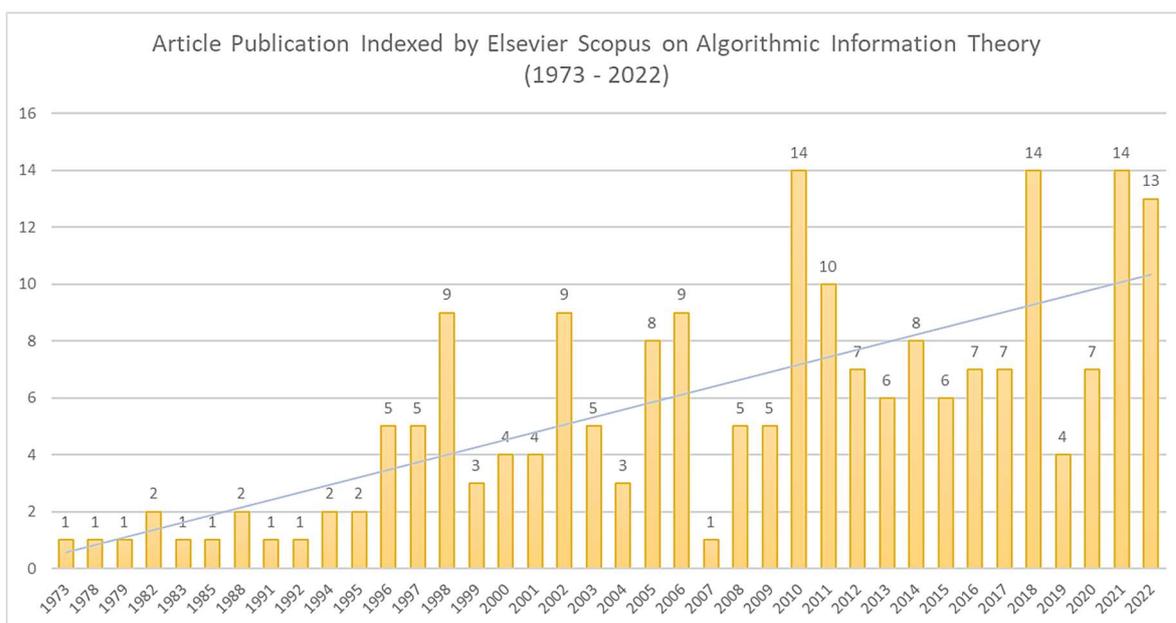


Gráfico 3: Distribuição da amostra em relação a ano de publicação.

A lista completa de artigos se encontra na Tabela 2, sendo que dos 207 apenas 131 foi, até o momento, possível obter o texto completo em PDF através do *Portal de Periódicos da CAPES* durante a elaboração deste trabalho.

Discussão

A AIT demonstra sua capacidade de englobar um conjunto poderoso de ideias com um amplo poder explicativo, abrangendo diversas áreas do conhecimento. Enquanto os artigos dos anos iniciais tendem a enfatizar o desenvolvimento teórico, à medida que avançamos ano após ano uma maior quantidade de publicações passa a se dedicar às aplicações. Por exemplo, em um estudo de 2022, foram aplicadas as ideias da AIT para analisar a similaridade entre o vírus SARS CoV-2 e outras cepas virais (CILIBRASI, VITÁNYI, 2022). Também em 2022, pesquisadores propuseram a utilização da AIT para aprimorar implementações de smart-homes (HOUZÉ, DESSALLES, *et al.*, 2022). Em 2018, a AIT foi empregada na análise da linguagem natural (DĘBOWSKI, 2018), enquanto em 2017 um estudo investigou a percepção da beleza e sua relação com a complexidade através dos princípios da AIT (GAUVRIT, SOLER-TOSCANO, *et al.*, 2017).

Diversos pesquisadores têm utilizado conceitos derivados da AIT para abordar o problema da consciência, como evidenciado em estudos de Maguire, Moser *et al.* (2016), Scheffel (2020) e Verheyen (2021). Além disso, a AIT tem sido aplicada no estudo da evolução biológica por pesquisadores como Agapow (2008), Ané, Sanderson *et al.* (2005), Corominas-Murtra, Seoane *et al.* (2018), Devine (2016), Johnston, Dingle *et al.* (2022), Miłkowski (2009), Oexle (1995) e Stebbing (2006). Esses estudos demonstram a versatilidade e a relevância da AIT em diferentes campos de pesquisa, contribuindo para uma melhor compreensão de fenômenos complexos e sua aplicação em diversas áreas acadêmicas.

Conclusão

Exploramos aqui a abrangência e a relevância da AIT em uma variedade de áreas de pesquisa. A partir da análise de estudos acadêmicos que empregaram os conceitos da AIT, podemos observar sua crescente aplicação em campos como análise de similaridade viral, aprimoramento de implementações tecnológicas, estudo da linguagem natural, percepção da beleza, consciência e evolução biológica. Esses exemplos demonstram o poder explanatório e a versatilidade da AIT como um referencial teórico que oferece explicações e abordagens inovadoras para compreender fenômenos complexos.

Além disso, a AIT desempenha um papel fundamental na busca por um entendimento mais abrangente da complexidade e da informação em diversas disciplinas acadêmicas. Seu enfoque na quantificação da complexidade e na análise de padrões e estruturas nos permite explorar as relações entre sistemas complexos, a natureza da informação e os limites da computação.

Com base nas evidências apresentadas, entendemos também que a AIT desempenha um papel crucial na condução de pesquisas interdisciplinares, multidisciplinares e transdisciplinares. Ademais, à medida que avançamos em direção a um mundo cada vez mais interconectado, a AIT se mostra uma ferramenta valiosa para compreender e descrever a complexidade subjacente a sistemas artificiais.

Futuros estudos são necessários para aprofundar os conhecimentos até aqui discutidos, principalmente através da leitura e análise do material obtido na etapa final.

Tabela 2: Lista de palavras mais usadas nos resumos

Freq. (%)	Termo	Freq. (%)	Termo	Freq. (%)	Termo
0,58	application	1,07	states	2,11	nature
0,58	artificial	1,07	theorem	2,40	set
0,58	halting	1,17	infinite	2,44	quantum
0,62	economic	1,20	intelligence	2,63	machine
0,65	energy	1,20	pattern	2,63	probability
0,65	enumerable	1,20	turing	2,66	entropy
0,65	prediction	1,27	real	2,73	bit
0,68	biological	1,33	chaitin	2,89	problem
0,75	mathematical	1,40	thermodynamic	2,89	process
0,78	cognitive	1,43	network	3,12	model
0,78	law	1,62	statistics	3,57	kolmogorov
0,81	evolution	1,75	theoretic	3,64	data
0,88	program	1,85	number	3,80	sequence
0,97	dynamics	1,88	finite	4,55	scientific
1,04	language	2,05	physics	4,71	computation
1,07	coding	2,08	function	4,84	randomness
		2,11	compression	10,65	complexity

Artigos obtidos no levantamento bibliográfico

- AARONSON, Scott. The Equivalence of Sampling and Searching. **Theory of Computing Systems**, v. 55, n. 2, p. 281-298-281-298, 2014.
- ABRAHÃO, Felipe S.; WEHMUTH, Klaus; ZENIL, Hector; *et al.* Algorithmic information distortions in node-aligned and node-unaligned multidimensional networks. **Entropy**, v. 23, n. 7, p. 835-835, 2021.
- AGAPOW, Paul-Michael. Building a better worm: Organisation and evolution in development. **Complexity International**, v. 12, p. 1-10-1-10, 2008.
- ANÉ, Cécile; SANDERSON, Michael J. Missing the forest for the trees: Phylogenetic compression and its implications for inferring complex evolutionary histories. **Systematic Biology**, v. 54, n. 1, p. 146-157-146-157, 2005.
- AY, Nihat; MÜLLER, Markus; SZKOŁA, Arleta. Effective complexity and its relation to logical depth. **IEEE Transactions on Information Theory**, v. 56, n. 9, p. 4593-4607-4593-4607, 2010.
- AY, Nihat; MÜLLER, Markus; SZKOŁA, Arleta. Effective complexity of stationary process realizations. **Entropy**, v. 13, n. 6, p. 1200-1211-1200-1211, 2011.
- BAEZ, John; STAY, Mike. Algorithmic thermodynamics. **Mathematical Structures in Computer Science**, v. 22, n. 5, p. 771-787-771-787, 2012.
- BARMPALIAS, George; CENZER, Douglas; PORTER, Christopher P. The probability of a computable output from a random oracle. **ACM Transactions on Computational Logic**, v. 18, n. 3, 2017.
- BARMPALIAS, George; LEWIS-PYE, Andrew. Compression of Data Streams Down to Their Information Content. **IEEE Transactions on Information Theory**, v. 65, n. 7, p. 4471-4485, 2019.
- BARMPALIAS, George; LEWIS-PYE, Andrew. Optimal redundancy in computations from random oracles. **Journal of Computer and System Sciences**, v. 92, p. 1-8-1-8, 2018.
- BARONCHELLI, Andrea; CAGLIOTI, Emanuele; LORETO, Vittorio. Measuring complexity with zippers. **European Journal of Physics**, v. 26, n. 5, p. S69-S77-S69-S77, 2005.
- BAUWENS, Bruno; GÁCS, Peter; ROMASHCHENKO, Andrei; *et al.* Inequalities for space-bounded Kolmogorov complexity. **Computability**, v. 11, n. 3-4, p. 165-185-165-185, 2022.
- BÉDARD, Charles Alexandre; BERGERON, Geoffroy. An Algorithmic Approach to Emergence. **Entropy**, v. 24, n. 7, 2022.
- BENNETT, Charles H; GÁCS, Péter; LI, Ming; *et al.* Information distance. **IEEE Transactions on Information Theory**, v. 44, n. 4, p. 1407-1423-1407-1423, 1998.
- BIENVENU, Laurent; DESFONTAINES, Damien; SHEN, Alexander. Generic algorithms for halting problem and optimal machines revisited. **Logical Methods in Computer Science**, v. 12, n. 2, 2016.
- BIESAGA, Mikołaj; TALAGA, Szymon; NOWAK, Andrzej. The Effect of Context and Individual Differences in Human-Generated Randomness. **Cognitive Science**, v. 45, n. 12, 2021.
- BONANNO, Claudio. A Complexity Approach to the Soliton Resolution Conjecture. **Journal of Statistical Physics**, v. 160, n. 5, p. 1432-1448-1432-1448, 2015.
- BONNICI, Vincenzo; MANCA, Vincenzo. an Informational Test for Random Finite Strings. **Entropy**, v. 20, n. 12, p. 934-934, 2018.
- BRANDOUY, Olivier; DELAHAYE, Jean-Paul; MA, Lin; *et al.* Algorithmic complexity of financial motions. **Research in International Business and Finance**, v. 30, p. 336-347-336-347, 2014.
- BRANDOUY, Olivier; DELAHAYE, Jean-Paul; MA, Lin. Estimating the algorithmic complexity of stock markets. **Algorithmic Finance**, v. 4, n. 3-4, p. 159-178-159-178, 2015.
- BRATTICO, Pauli. Shallow Reductionism and the Problem of Complexity in Psychology. **Theory & Psychology**, v. 18, n. 4, p. 483-504-483-504, 2008.
- BROWN, Daniel G.; MONDOL, Tiasa. On the problem of small objects. **Entropy**, v. 23, n. 11, p. 1524-1524, 2021.

- CALUDE, Cristian. Algorithmic information theory: open problems 1. **Journal of Universal Computer Science**, v. 2, n. 5, p. 439-441-439-441, 1996.
- CALUDE, Cristian; GROZEA, Cristian. Kraft-Chaitin inequality revisited. **Journal of Universal Computer Science**, v. 2, n. 5, p. 306-310-306-310, 1996.
- CALUDE, Cristian S; DINNEEN, Michael J; DUMITRESCU, Monica; *et al.* Experimental evidence of quantum randomness incomputability. **Physical Review A - Atomic, Molecular, and Optical Physics**, v. 82, n. 2, 2010.
- CALUDE, Cristian S.; DUMITRESCU, Monica. Entropic measures, Markov information sources and complexity. **Applied Mathematics and Computation**, v. 132, n. 2-3, p. 369-384, 2002.
- CALUDE, Cristian S.; JAIN, Sanjay; MERKLE, Wolfgang; *et al.* Searching for shortest and least programs. **Theoretical Computer Science**, v. 807, p. 114-127, 2020.
- CALUDE, Cristian S; LONGO, Giuseppe. The Deluge of Spurious Correlations in Big Data. **Foundations of Science**, v. 22, n. 3, p. 595-612-595-612, 2017.
- CALUDE, Cristian S; NIES, André; STAIGER, Ludwig; *et al.* Universal recursively enumerable sets of strings. **Theoretical Computer Science**, v. 412, n. 22, p. 2253-2261-2253-2261, 2011.
- CALUDE, Cristian S; SALOMAA, Kai; ROBLLOT, Tania K. Finite state complexity. **Theoretical Computer Science**, v. 412, n. 41, p. 5668-5677-5668-5677, 2011.
- CALUDE, Cristian S; SALOMAA, Kai; ROBLLOT, Tania K. State-size hierarchy for finite-state complexity. **International Journal of Foundations of Computer Science**, v. 23, n. 1, p. 37-50-37-50, 2012.
- CALUDE, Cristian S; STAIGER, Ludwig; SVOZIL, Karl. Randomness relative to Cantor expansions. **Communications in Nonlinear Science and Numerical Simulation**, v. 10, n. 8, p. 921-930-921-930, 2005.
- CALUDE, Cristian S; ZIMAND, Marius. Algorithmically independent sequences. **Information and Computation**, v. 208, n. 3, p. 292-308-292-308, 2010.
- CÂMPEANU, Cezar. Descriptive complexity in Encoded Blum Static Complexity spaces. **International Journal of Foundations of Computer Science**, v. 25, n. 7, p. 917-932-917-932, 2014.
- CASE, Adam; LUTZ, Jack H. Mutual dimension and random sequences. **Theoretical Computer Science**, v. 731, p. 68-87-68-87, 2018.
- CERRA, Daniele; DATCU, Mihai. Algorithmic relative complexity. **Entropy**, v. 13, n. 4, p. 902-914, 2011.
- CERRA, Daniele; DATCU, Mihai. Expanding the algorithmic information theory frame for applications to earth observation. **Entropy**, v. 15, n. 1, p. 407-415, 2013.
- CERRA, Daniele; MALLET, Alexandre; GUEGUEN, Lionel; *et al.* Algorithmic information theory-based analysis of earth observation images: An assessment. **IEEE Geoscience and Remote Sensing Letters**, v. 7, n. 1, p. 8-12-8-12, 2010.
- CHAITIN, G J. A new version of algorithmic information theory. **Complexity**, v. 1, n. 4, p. 55-59-55-59, 1996.
- CHAITIN, Gregory. Two philosophical applications of algorithmic information theory. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 2731, p. 1-10-1-10, 2003.
- CHAITIN, Gregory J. How to run algorithmic information theory on a computer: Studying the limits of mathematical reasoning. **Complexity**, v. 2, n. 1, p. 15-21-15-21, 1996.
- CHAITIN, Gregory J; SCHWARTZ, Jacob T. A note on monte carlo primality tests and algorithmic information theory. **Communications on Pure and Applied Mathematics**, v. 31, n. 4, p. 521-527-521-527, 1978.
- CHEN, Shu-Heng; YEH, Chia-Hsuan. Simulating economic transition processes by genetic programming. **Annals of Operations Research**, v. 97, n. 1-4, p. 265-286-265-286, 2000.

- CILIBRASI, Rudi L.; VITÁNYI, Paul M.B. Fast Phylogeny of SARS-CoV-2 by Compression. **Entropy**, v. 24, n. 4, p. 439–439, 2022.
- COHEN, Andrew R; BJORNSSON, Christopher S; TEMPLE, Sally; *et al.* Automatic summarization of changes in biological image sequences using algorithmic information theory. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 31, n. 8, p. 1386-1403-1386–1403, 2009.
- COHEN, Andrew R; GOMES, Francisco L A F; ROYSAM, Badrinath; *et al.* Computational prediction of neural progenitor cell fates. **Nature Methods**, v. 7, n. 3, p. 213-218-213–218, 2010.
- CONNOR, Richard; SIMEONI, Fabio; IAKOVOS, Michael; *et al.* A bounded distance metric for comparing tree structure. **Information Systems**, v. 36, n. 4, p. 748–764, 2011.
- COROMINAS-MURTRA, Bernat; SEOANE, Luís F.; SOLÉ, Ricard. Zipf's Law, unbounded complexity and open-ended evolution. **Journal of the Royal Society Interface**, v. 15, n. 149, p. 20180395–20180395, 2018.
- COROMINAS-MURTRA, Bernat; SOLÉ, Ricard V. Universality of Zipf's law. **Physical Review E - Statistical, Nonlinear, and Soft Matter Physics**, v. 82, n. 1, 2010.
- CRUTCHFIELD, J P; PACKARD, N H. Symbolic dynamics of one-dimensional maps: Entropies, finite precision, and noise. **International Journal of Theoretical Physics**, v. 21, n. 6–7, p. 433-466-433–466, 1982.
- DAY, Adam R. Increasing the gap between descriptive complexity and algorithmic probability. **Transactions of the American Mathematical Society**, v. 363, n. 10, p. 5577-5604-5577–5604, 2011.
- DAY, Adam R. On the computational power of random strings. **Annals of Pure and Applied Logic**, v. 160, n. 2, p. 214-228-214–228, 2009.
- DAY, Adam R. The computable Lipschitz degrees of computably enumerable sets are not dense. **Annals of Pure and Applied Logic**, v. 161, n. 12, p. 1588-1602-1588–1602, 2010.
- DĘBOWSKI, Łukasz. Is natural language a perigraphic process? The theorem about facts and words revisited. **Entropy**, v. 20, n. 2, 2018.
- DEVINE, S. The information requirements of complex biological and economic systems with algorithmic information theory. **International Journal of Design and Nature and Ecodynamics**, v. 12, n. 3, p. 367–376, 2017.
- DEVINE, Sean. Algorithmic entropy and landauer's principle link microscopic system behaviour to the thermodynamic entropy. **Entropy**, v. 20, n. 10, p. 798–798, 2018.
- DEVINE, Sean. An algorithmic information theory challenge to intelligent design. **Zygon**, v. 49, n. 1, p. 42–65, 2014.
- DEVINE, Sean. An economy viewed as a far-from-equilibrium system from the perspective of algorithmic information theory. **Entropy**, v. 20, n. 4, p. 228–228, 2018.
- DEVINE, Sean. Landauer's principle a consequence of bit flows, given stirling's approximation. **Entropy**, v. 23, n. 10, p. 1288–1288, 2021.
- DEVINE, Sean. The application of algorithmic information theory to noisy patterned strings. **Complexity**, v. 12, n. 2, p. 52–58, 2006.
- DEVINE, Sean D. Understanding how replication processes can maintain systems away from equilibrium using Algorithmic Information Theory. **BioSystems**, v. 140, p. 8–22, 2016.
- DINGLE, Kamaludin; CAMARGO, Chico Q; LOUIS, Ard A. Input-output maps are strongly biased towards simple outputs. **Nature Communications**, v. 9, n. 1, 2018.
- DINGLE, Kamaludin; NOVEV, Javor K; AHNERT, Sebastian E; *et al.* Predicting phenotype transition probabilities via conditional algorithmic probability approximations. **Journal of the Royal Society Interface**, v. 19, n. 197, 2022.
- DINGLE, Kamaludin; PÉREZ, Guillermo Valle; LOUIS, Ard A. Generic predictions of output probability based on complexities of inputs and outputs. **Scientific Reports**, v. 10, n. 1, p. 4415–4415, 2020.

- DOWNEY, Rod G; HIRSCHFELDT, Denis R; LAFORTE, Geoff. Randomness and reducibility. **Journal of Computer and System Sciences**, v. 68, n. 1, p. 96-114-96-114, 2004.
- DOWNEY, Rod G; HIRSCHFELDT, Denis R; NIES, André. Randomness, computability, and density. **SIAM Journal on Computing**, v. 31, n. 4, p. 1169-1183-1169-1183, 2002.
- DOWNEY, Rodney G; LAFORTE, Geoffrey L. Presentations of computably enumerable reals. **Theoretical Computer Science**, v. 284, n. 2, p. 539-555-539-555, 2002.
- EPSTEIN, Samuel. An extended coding theorem with application to quantum complexities. **Information and Computation**, v. 275, 2020.
- FERBUS-ZANDA, Marie; GRIGORIEFF, Serge. Kolmogorov complexity and set theoretical representations of integers. **Mathematical Logic Quarterly**, v. 52, n. 4, p. 375-403-375-403, 2006.
- GÁCS, Péter. On the relation between descriptonal complexity and algorithmic probability. **Theoretical Computer Science**, v. 22, n. 1-2, p. 71-93-71-93, 1983.
- GÁCS, Peter. Quantum algorithmic entropy. **Journal of Physics A: Mathematical and General**, v. 34, n. 35, p. 6859-6880-6859-6880, 2001.
- GÁCS, Peter. Uniform test of algorithmic randomness over a general space. **Theoretical Computer Science**, v. 341, n. 1-3, p. 91-137, 2005.
- GÁCS, Péter; TROMP, John T; VITÁNYI, Paul M B. Algorithmic statistics. **IEEE Transactions on Information Theory**, v. 47, n. 6, p. 2443-2463-2443-2463, 2001.
- GAUVRIT, Nicolas; SOLER-TOSCANO, Fernando; GUIDA, Alessandro. A preference for some types of complexity comment on “perceived beauty of random texture patterns: A preference for complexity”. **Acta Psychologica**, v. 174, p. 48-53-48-53, 2017.
- GOERTZEL, Ben. Self-organizing evolution. **Journal of Social and Evolutionary Systems**, v. 15, n. 1, p. 7-53-7-53, 1992.
- GORBUNOV, K.Yu. On a complexity of the formula $(A \vee B) \rightarrow C$. **Theoretical Computer Science**, v. 207, n. 2, p. 383-386-383-386, 1998.
- GRANADOS, Ana; MARTÍNEZ, Rafael; CAMACHO, David; *et al.* Improving NCD accuracy by combining document segmentation and document distortion. **Knowledge and Information Systems**, v. 41, n. 1, p. 223-245-223-245, 2014.
- GUREVICH, Yuri; PASSMORE, Grant Olney. Impugning Randomness, Convincingly. **Studia Logica**, v. 100, n. 1-2, p. 193-222-193-222, 2012.
- HĂLĂLAE, Ioan. Cellular automata with dynamical loop function as noise generators: Statistical aspects (1). **UPB Scientific Bulletin, Series C: Electrical Engineering**, v. 72, n. 3, p. 73-82-73-82, 2010.
- HAY, Nicholas J. Simulation complexity. **Fundamenta Informaticae**, v. 83, n. 1-2, p. 117-140-117-140, 2008.
- HERNÁNDEZ-ORALLO, José. Evaluation in artificial intelligence: from task-oriented to ability-oriented measurement. **Artificial Intelligence Review**, v. 48, n. 3, p. 397-447-397-447, 2017.
- HERNÁNDEZ-ORALLO, José. On environment difficulty and discriminating power. **Autonomous Agents and Multi-Agent Systems**, v. 29, n. 3, p. 402-454-402-454, 2015.
- HERNÁNDEZ-ORALLO, José; DOWE, David L. Measuring universal intelligence: Towards an anytime intelligence test. **Artificial Intelligence**, v. 174, n. 18, p. 1508-1539-1508-1539, 2010.
- HERNÁNDEZ-ORALLO, José; DOWE, David L; HERNÁNDEZ-LLOREDA, M Victoria. Universal psychometrics: Measuring cognitive abilities in the machine kingdom. **Cognitive Systems Research**, v. 27, p. 50-74-50-74, 2014.
- HERTLING, Peter. Nonrandom sequences between random sequences. **Journal of Universal Computer Science**, v. 11, n. 12, p. 1970-1985-1970-1985, 2005.
- HITCHCOCK, John M; LUTZ, Jack H. Why computational complexity requires stricter martingales. **Theory of Computing Systems**, v. 39, n. 2, p. 277-296-277-296, 2006.

- HITCHCOCK, John M.; MAYORDOMO, Elvira. Base invariance of feasible dimension. **Information Processing Letters**, v. 113, n. 14–16, p. 546–551, 2013.
- HOFFMANN, Achim G. Complex connectionist dynamics and the limited complexity of resulting functions. **Connection Science**, v. 9, n. 2, p. 201-216-201–216, 1997.
- HOFFMANN, Achim G. On computational limitations of self-organization. **Neurocomputing**, v. 15, n. 1, p. 69-87-69–87, 1997.
- HOUZÉ, Étienne; DESSALLES, Jean Louis; DIACONESCU, Ada; *et al.* What Should I Notice? Using Algorithmic Information Theory to Evaluate the Memorability of Events in Smart Homes. **Entropy**, v. 24, n. 3, p. 346–346, 2022.
- HUTTER, Marcus. On generalized computable universal priors and their convergence. **Theoretical Computer Science**, v. 364, n. 1, p. 27-41-27–41, 2006.
- HUTTER, Marcus. Sequential predictions based on algorithmic complexity. **Journal of Computer and System Sciences**, v. 72, n. 1, p. 95-117-95–117, 2006.
- HUTTER, Marcus; MUCHNIK, Andrej. On semimeasures predicting Martin-Löf random sequences. **Theoretical Computer Science**, v. 382, n. 3, p. 247-261-247–261, 2007.
- JALALI, Shirin; MALEKI, Arian; BARANIUK, Richard G. Minimum complexity pursuit for universal compressed sensing. **IEEE Transactions on Information Theory**, v. 60, n. 4, p. 2253-2268-2253–2268, 2014.
- JANZING, Dominik; STEUDEL, Bastian. Justifying additive noise model-based causal discovery via algorithmic information theory. **Open Systems and Information Dynamics**, v. 17, n. 2, p. 189-212-189–212, 2010.
- JIAO, Rong; LIN, Nan; HU, Zixin; *et al.* Bivariate causal discovery and its applications to gene expression and imaging data analysis. **Frontiers in Genetics**, v. 9, n. AUG, 2018.
- JOHNSTON, Iain G.; DINGLE, Kamaludin; GREENBURY, Sam F.; *et al.* Symmetry and simplicity spontaneously emerge from the algorithmic nature of evolution. **Proceedings of the National Academy of Sciences of the United States of America**, v. 119, n. 11, p. e2113883119–e2113883119, 2022.
- KOLCHINSKY, Artemy. A Novel Approach to the Partial Information Decomposition. **Entropy**, v. 24, n. 3, 2022.
- KOLCHINSKY, Artemy; WOLPERT, David H. Thermodynamic costs of Turing machines. **Physical Review Research**, v. 2, n. 3, 2020.
- KOPPL, Roger. Some epistemological implications of economic complexity. **Journal of Economic Behavior and Organization**, v. 76, n. 3, p. 859-872-859–872, 2010.
- KOSHELEV, Misha; LONGPRÉ, Luc. Algorithmic information theory may explain the pathogenic number of DNA repeats in myotonic dystrophy type 1 (And in similar diseases). **Journal of Uncertain Systems**, v. 9, n. 2, p. 113-123-113–123, 2015.
- KOTSIREAS, Ilias S; KOUKOUVINOS, Christos; PARDALOS, Panos M. An efficient string sorting algorithm for weighing matrices of small weight. **Optimization Letters**, v. 4, n. 1, p. 29-36-29–36, 2010.
- KOZACHINSKIY, Alexander; SHEN, Alexander. Automatic Kolmogorov complexity, normality, and finite-state dimension revisited. **Journal of Computer and System Sciences**, v. 118, p. 75–107, 2021.
- KRASNOGOR, Natalio; PELTA, D A. Measuring the similarity of protein structures by means of the Universal Similarity Metric. **Bioinformatics**, v. 20, n. 7, p. 1015-1021-1015–1021, 2004.
- KREINOVICH, Vladik; LONGPRÉ, Luc. Pure quantum states are fundamental, mixtures (composite states) are mathematical constructions: An argument using algorithmic information theory. **International Journal of Theoretical Physics**, v. 36, n. 1, p. 167-176-167–176, 1997.
- KREINOVICH, Vladik; LONGPRÉ, Luc. Why Kolmogorov complexity in physical equations? **International Journal of Theoretical Physics**, v. 37, n. 11, p. 2791-2801-2791–2801, 1998.
- KUDIN, A M. Cryptographic transformations of non-shannon sources of information. **Cybernetics and Systems Analysis**, v. 46, n. 5, p. 813-819-813–819, 2010.

- KURTZ, Stuart A. A computer scientist's perspective on chaos and mystery. **Zygon**, v. 37, n. 2, p. 415-420-415-420, 2002.
- LEMUS, Vladimir; ACUÑA-YEOMANS, Eduardo; ZAMORA, Víctor; *et al.* Computable model discovery and high-level-programming approximations to algorithmic complexity. **Theoretical Computer Science**, v. 927, p. 31-41, 2022.
- LI, Ming; VITÁNYI, Paul M B. Tape versus queue and stacks: The lower bounds. **Information and Computation**, v. 78, n. 1, p. 56-85-56-85, 1988.
- LIMA, Matheus Sant Ana. Information theory inspired optimization algorithm for efficient service orchestration in distributed systems. **PLoS one**, v. 16, n. 1, p. e0242285-e0242285, 2021.
- LIU, Jing; PAN, Zhiwen; XU, Jingce; *et al.* Quality-time-complexity universal intelligence measurement. **International Journal of Crowd Science**, v. 2, n. 2, p. 99-107, 2018.
- LLOYD, Seth; ZUREK, Wojciech H. Algorithmic treatment of the spin-echo effect. **Journal of Statistical Physics**, v. 62, n. 3-4, p. 819-839, 1991.
- LONGPRÉ, Luc; KREINOVICH, Vladik. When are two wave functions distinguishable: A new answer to Pauli's question, with potential application to quantum cosmology. **International Journal of Theoretical Physics**, v. 47, n. 3, p. 814-831-814-831, 2008.
- LUTZ, Jack H. Effective fractal dimensions. **Mathematical Logic Quarterly**, v. 51, n. 1, p. 62-72, 2005.
- MACHTA, J. Entropy, information, and computation. **American Journal of Physics**, v. 67, n. 12, p. 1074-1077-1074-1077, 1999.
- MACIEJOWSKI, J M. Model discrimination using an algorithmic information criterion. **Automatica**, v. 15, n. 5, p. 579-593-579-593, 1979.
- MAGUIRE, Phil; MOSER, Philippe; MAGUIRE, Rebecca; *et al.* Seeing Patterns in Randomness: A Computational Model of Surprise. **Topics in Cognitive Science**, v. 11, n. 1, p. 103-118, 2019.
- MAGUIRE, Phil; MOSER, Philippe; MAGUIRE, Rebecca. Understanding consciousness as data compression. **Journal of Cognitive Science**, v. 17, n. 1, p. 63-94-63-94, 2016.
- MARKOWSKY, George. An introduction to algorithmic information theory. **Complexity**, v. 2, n. 4, p. 14-22-14-22, 1997.
- MCALLISTER, James W. Algorithmic randomness in empirical data. **Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences**, v. 34, n. 3, p. 633-646, 2003.
- MÉNDEZ, I. Model selection for radiochromic film dosimetry. **Physics in Medicine and Biology**, v. 60, n. 10, p. 4089-4104-4089-4104, 2015.
- MIŁKOWSKI, Marcin. Is evolution algorithmic? **Minds and Machines**, v. 19, n. 4, p. 465-475, 2009.
- MILOSAVLJEVIĆ, A. Sequence comparisons via algorithmic mutual information. **Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology**, v. 2, p. 303-309-303-309, 1994.
- MILOSAVLJEVIĆ, Aleksandar. Discovering Dependencies via Algorithmic Mutual Information: A Case Study in DNA Sequence Comparisons. **Machine Learning**, v. 21, n. 1, p. 35-50-35-50, 1995.
- MILOVANOV, Alexey. Some Properties of Antistochastic Strings. **Theory of Computing Systems**, v. 61, n. 2, p. 521-535-521-535, 2017.
- MIYADERA, Takayuki; IMAI, Hideki. Quantum Kolmogorov complexity and quantum key distribution. **Physical Review A - Atomic, Molecular, and Optical Physics**, v. 79, n. 1, 2009.
- MONDOL, Tiasa; BROWN, Daniel G. Computational creativity and aesthetics with algorithmic information theory. **Entropy**, v. 23, n. 12, p. 1654-1654, 2021.
- MOSER, Philippe. Polylog depth, highness and lowness for E. **Information and Computation**, v. 271, 2020.
- MUCHNIK, An A. On common information. **Theoretical Computer Science**, v. 207, n. 2, p. 319-328, 1998.

- MUCHNIK, Andrej A. Conditional complexity and codes. **Theoretical Computer Science**, v. 271, n. 1–2, p. 97-109-97–109, 2002.
- MUCHNIK, Andrej A. Kolmogorov complexity and cryptography. **Proceedings of the Steklov Institute of Mathematics**, v. 274, n. 1, p. 193-203-193–203, 2011.
- NABUTOVSKY, Alexander; WEINBERGER, Shmuel. The fractal nature of Riem/Diff I. **Geometriae Dedicata**, v. 101, n. 1, p. 1-54-1–54, 2003.
- NANDAKUMAR, Satyadev; VISHNOI, Prateek. On continued fraction randomness and normality. **Information and Computation**, v. 285, 2022.
- NAYYERI, Amirahmad; DASTGHAIBYFARD, Gholamhossein. An information theoretic analysis of random number generator based on cellular automaton. **International Journal of Advanced Computer Science and Applications**, v. 9, n. 1, p. 321-329-321–329, 2018.
- NEFEDIEV, L A; SAKHBIEVA, A R. The information coding in the time structure of the object of a laser pulse in an optical echo processor. **Advances in Optical Technologies**, 2012.
- NYKTER, Matti; PRICE, Nathan D.; ALDANA, Maximino; *et al.* Gene expression dynamics in the macrophage exhibit criticality. **Proceedings of the National Academy of Sciences of the United States of America**, v. 105, n. 6, p. 1897–1900, 2008.
- OEXLE, Konrad. Data compressibility, physical entropy, and evolutionary a priori relation between observer and object. **Physical Review E**, v. 51, n. 3, p. 2651-2653-2651–2653, 1995.
- ORD, Toby; KIEU, Tien D. On the existence of a new family of Diophantine equations for Ω . **Fundamenta Informaticae**, v. 56, n. 3, p. 273-284-273–284, 2003.
- ORPONEN, Tuomas. Combinatorial proofs of two theorems of Lutz and Stull. **Mathematical Proceedings of the Cambridge Philosophical Society**, v. 171, n. 3, p. 503-514-503–514, 2021.
- PAJUNEN, Petteri. Blind source separation using algorithmic information theory. **Neurocomputing**, v. 22, n. 1–3, p. 35-48-35–48, 1998.
- PANTAZI, Stefan V; KUSHNIRUK, Andre; MOEHR, Jochen R. The usability axiom of medical information systems. **International Journal of Medical Informatics**, v. 75, n. 12, p. 829-839-829–839, 2006.
- PEPTENATU, Daniel; ANDRONACHE, Ion; AHAMMER, Helmut; *et al.* Kolmogorov compression complexity may differentiate different schools of Orthodox iconography. **Scientific Reports**, v. 12, n. 1, p. 10743–10743, 2022.
- PIANTADOSI, Steven T; FEDORENKO, Evelina. Infinitely productive language can arise from chance under communicative pressure. **Journal of Language Evolution**, v. 2, n. 2, p. 141-147-141–147, 2017.
- POLAND, Jan. A coding theorem for Enumerable Output Machines. **Information Processing Letters**, v. 91, n. 4, p. 157-161-157–161, 2004.
- POLAND, Jan; HUTTER, Marcus. Asymptotics of discrete MDL for online prediction. **IEEE Transactions on Information Theory**, v. 51, n. 11, p. 3780-3795-3780–3795, 2005.
- PORTER, Christopher P. Revisiting chaitin’s incompleteness theorem. **Notre Dame Journal of Formal Logic**, v. 62, n. 1, p. 147-171-147–171, 2021.
- RAATIKAINEN, Panu. Algorithmic information theory and undecidability. **Synthese**, v. 123, n. 2, p. 217-225-217–225, 2000.
- RAATIKAINEN, Panu. On interpreting Chaitin’s incompleteness theorem. **Journal of Philosophical Logic**, v. 27, n. 6, p. 569-586-569–586, 1998.
- RATHMANNER, Samuel; HUTTER, Marcus. A philosophical treatise of universal induction. **Entropy**, v. 13, n. 6, p. 1076-1136-1076–1136, 2011.
- RAVARI, Alireza Norouzzadeh; TAGHIRAD, Hamid D. Loop closure detection by algorithmic information theory: Implemented on range and camera image data. **IEEE Transactions on Cybernetics**, v. 44, n. 10, p. 1938-1949-1938–1949, 2014.
- RESENDE, João S.; MARTINS, Rolando; ANTUNES, Luís. A survey on using Kolmogorov Complexity in cybersecurity. **Entropy**, v. 21, n. 12, p. 1196–1196, 2019.

- RICCI, Gabriele. Syntactic Compression Codes at the Zero Entropy Point. **International Journal of Computer Mathematics**, v. 11, n. 3–4, p. 221–226, 1982.
- ROBERTSON, Douglas S. Algorithmic information theory, free will, and the Turing test. **Complexity**, v. 4, n. 3, p. 25–34–25–34, 1999.
- ROMANO, Sergio; SIGMAN, Mariano; FIGUEIRA, Santiago. LT 2 C 2 : A language of thought with Turing-computable Kolmogorov complexity. **Papers in Physics**, v. 5, 2013.
- ROMASHCHENKO, A; SHEN, A; VERESHCHAGIN, N. Combinatorial interpretation of Kolmogorov complexity. **Proceedings of the Annual IEEE Conference on Computational Complexity**, p. 131–137–131–137, 2000.
- ROMASHCHENKO, A.; SHEN, A.; VERESHCHAGIN, N. Combinatorial interpretation of Kolmogorov complexity. **Theoretical Computer Science**, v. 271, n. 1–2, p. 111–123, 2002.
- RUFFINI, Giulio; IBAÑEZ, David; KROUPI, Eleni; *et al.* Algorithmic Complexity of EEG for Prognosis of Neurodegeneration in Idiopathic Rapid Eye Movement Behavior Disorder (RBD). **Annals of Biomedical Engineering**, v. 47, n. 1, p. 282–296, 2019.
- RYABKO, Boris; ASTOLA, Jaakko; GAMMERMAN, Alex. Application of Kolmogorov complexity and universal codes to identity testing and nonparametric testing of serial independence for time series. **Theoretical Computer Science**, v. 359, n. 1–3, p. 440–448–440–448, 2006.
- SADOT, Avital; SARBU, Septimia; KESSELI, Juha; *et al.* Information-Theoretic Analysis of the Dynamics of an Executable Biological Model. **PLoS ONE**, v. 8, n. 3, 2013. >.
- SANTORO, Andrea; NICOSIA, Vincenzo. Algorithmic Complexity of Multiplex Networks. **Physical Review X**, v. 10, n. 2, p. 21069–21069, 2020.
- SARKAR, Aritra; AL-ARS, Zaid; BERTELS, Koen. Estimating algorithmic information using quantum computing for genomics applications. **Applied Sciences (Switzerland)**, v. 11, n. 6, p. 2696–2696, 2021.
- SCHACK, Rüdiger. Algorithmic information and simplicity in statistical physics. **International Journal of Theoretical Physics**, v. 36, n. 1, p. 209–226, 1997.
- SCHEFFEL, Jan. On the Solvability of the Mind–Body Problem. **Axiomathes**, v. 30, n. 3, p. 289–312–289–312, 2020.
- SHEN, Alexander. Discussion on Kolmogorov complexity and statistical analysis. **Computer Journal**, v. 42, n. 4, p. 340–342–340–342, 1999.
- SHEN, Alexander. Individual codewords. **Theoretical Computer Science**, v. 908, p. 120–122–120–122, 2022.
- SHEN, Alexander; VERESHCHAGIN, Nikolai. Logical operations and Kolmogorov complexity. **Theoretical Computer Science**, v. 271, n. 1–2, p. 125–129–125–129, 2002.
- SIEBES, Arno. Data science as a language: challenges for computer science—a position paper. **International Journal of Data Science and Analytics**, v. 6, n. 3, p. 177–187–177–187, 2018.
- SILVA, Jorge Miguel; PRATAS, Diogo; ANTUNES, Rui; *et al.* Automatic analysis of artistic paintings using information-based measures. **Pattern Recognition**, v. 114, 2021.
- SILVA, Jorge Miguel; PRATAS, Diogo; CAETANO, Tânia; *et al.* The complexity landscape of viral genomes. **GigaScience**, v. 11, 2022.
- SOKLAKOV, Andrei N; SCHACK, Rüdiger. Preparation information and optimal decompositions for mixed quantum states. **Journal of Modern Optics**, v. 47, n. 12, p. 2265–2276–2265–2276, 2000.
- SORENSEN, Roy A. Yablo’s paradox and kindred infinite liars. **Mind**, v. 107, n. 425, p. 137–156–137–156, 1998.
- STAY, Michael. Very simple chaitin machines for concrete AIT. **Fundamenta Informaticae**, v. 68, n. 3, p. 231–247–231–247, 2005.
- STEBBING, A. R.D. Genetic parsimony: A factor in the evolution of complexity, order and emergence. **Biological Journal of the Linnean Society**, Received 3. v. 88, n. 2, p. 295–308, 2006.

- STERKENBURG, Tom F. Solomonoff prediction and Occam's razor. **Philosophy of Science**, v. 83, n. 4, p. 459-479-459-479, 2016.
- SUÑÉ, Abel; MARTÍNEZ, Manolo. Real patterns and indispensability. **Synthese**, v. 198, n. 5, p. 4315-4330, 2021.
- SVOZIL, K. Quantum algorithmic information theory. **Journal of Universal Computer Science**, v. 2, n. 5, p. 311-346-311-346, 1996.
- TADAKI, K. Fixed point theorems on partial randomness. **Annals of Pure and Applied Logic**, v. 163, n. 7, p. 763-774-763-774, 2012.
- TADAKI, Kohtaro. A generalization of Chaitin's halting probability Ω and halting self-similar sets. **Hokkaido Mathematical Journal**, v. 31, n. 1, p. 219-253-219-253, 2002.
- TADAKI, Kohtaro. A statistical mechanical interpretation of algorithmic information theory III: Composite systems and fixed points. **Mathematical Structures in Computer Science**, v. 22, n. 5, p. 752-770-752-770, 2012.
- TADAKI, Kohtaro. An extension of Chaitin's halting probability Ω to a measurement operator in an infinite dimensional quantum system. **Mathematical Logic Quarterly**, v. 52, n. 5, p. 419-438-419-438, 2006.
- TAKAHASHI, Hayato; AIHARA, Kazuyuki. Algorithmic analysis of irrational rotations in a single neuron model. **Journal of Complexity**, v. 19, n. 2, p. 132-152-132-152, 2003.
- TARTARA, Guido. On the application of algorithmic information theory to decision problems. **Information Sciences**, v. 6, n. C, p. 85-96, 1973.
- USPENSKY, V A; V'YUGIN, V V. Development of the algorithmic information theory in Russia. **Journal of Communications Technology and Electronics**, v. 56, n. 6, p. 739-747-739-747, 2011.
- VALLÉE, B. Dynamical sources in information theory: Fundamental intervals and word prefixes. **Algorithmica (New York)**, v. 29, n. 1-2, p. 262-306, 2001.
- VERHEYEN, Peter. From Information and Quantum Physics to Consciousness and Reality. **Sci**, v. 3, n. 4, p. 35-35, 2021.
- VIDAL, Clément. Computational and Biological Analogies for Understanding Fine-Tuned Parameters in Physics. **Foundations of Science**, v. 15, n. 4, p. 375-393-375-393, 2010.
- VITÁNYI, Paul M B. Quantum Kolmogorov complexity based on classical descriptions. **IEEE Transactions on Information Theory**, v. 47, n. 6, p. 2464-2479-2464-2479, 2001.
- WANG, Yongge. A comparison of two approaches to pseudorandomness. **Theoretical Computer Science**, v. 276, n. 1-2, p. 449-459-449-459, 2002.
- WEIJS, S. V.; VAN DE GIESEN, N.; PARLANGE, M. B. Data compression to define information content of hydrological time series. **Hydrology and Earth System Sciences**, v. 17, n. 8, p. 3171-3187, 2013.
- WEIJS, Steven V.; VAN DE GIESEN, Nick; PARLANGE, Marc B. HydroZIP: How hydrological knowledge can be used to improve compression of hydrological data. **Entropy**, v. 15, n. 4, p. 1289-1310, 2013.
- WHEELER, Billy. How realist is informational structural realism? **Synthese**, v. 200, n. 6, 2022.
- WHEELER, Billy. Simplicity, language-dependency and the best system account of laws. **Theoria (Spain)**, v. 31, n. 2, p. 189-206-189-206, 2016.
- WHITE, Ian. The Limits and Capabilities of Machines—A Review. **IEEE Transactions on Systems, Man and Cybernetics**, v. 18, n. 6, p. 917-938-917-938, 1988.
- WOLFF, J Gerard. Computing and information compression: A reply. **AI Communications**, v. 7, n. 3-4, p. 203-219-203-219, 1994.
- ZENIL, Hector; DELAHAYE, Jean-Paul. An algorithmic information theoretic approach to the behaviour of financial markets. **Journal of Economic Surveys**, v. 25, n. 3, p. 431-463-431-463, 2011.

ZENIL, Hector; KIANI, Narsis A.; SHANG, Ming Mei; *et al.* Algorithmic Complexity and Reprogrammability of Chemical Structure Networks. **Parallel Processing Letters**, v. 28, n. 1, p. 1850005–1850005, 2018.

ZENIL, Hector; KIANI, Narsis A; TEGNÉR, Jesper. Quantifying loss of information in network-based dimensionality reduction techniques. **Journal of Complex Networks**, v. 4, n. 3, p. 342-362-342–362, 2016.

ZENIL, Hector; SOLER-TOSCANO, Fernando; DELAHAYE, Jean Paul; *et al.* Two-dimensional Kolmogorov complexity and an empirical validation of the Coding theorem method by compressibility. **PeerJ Computer Science**, v. 2015, n. 9, p. e23–e23, 2015.

ZORBA, Serkan. God is random a novel argument for the existence of god. **European Journal of Science and Theology**, v. 12, n. 1, p. 51-67-51–67, 2016.